

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
FACULDADE DE DIREITO
DEPARTAMENTO DE CIÊNCIAS PENAIS

Vinicius de Fraga Bassotto

**RESPONSABILIZAÇÃO PENAL PELA DIVULGAÇÃO DE *FAKE NEWS* E POR
CRIMES DE ÓDIO PRATICADOS POR SISTEMAS DE INTELIGÊNCIA ARTIFICIAL
GENERATIVA**

PORTO ALEGRE
2023

VINICIUS DE FRAGA BASSOTTO

**RESPONSABILIZAÇÃO PENAL PELA DIVULGAÇÃO DE *FAKE NEWS* E POR
CRIMES DE ÓDIO PRATICADOS POR SISTEMAS DE INTELIGÊNCIA ARTIFICIAL
GENERATIVA**

Trabalho de Conclusão de Curso
apresentado como requisito parcial para
obtenção do grau de Bacharel em Direito
pela Faculdade de Direito da Universidade
Federal do Rio Grande do Sul.

Orientadora: Prof.^a Dr.^a Vanessa Chiari
Gonçalves

Porto Alegre
2023

CIP – Catalogação na Publicação

Bassotto, Vinicius de Fraga
Responsabilização penal pela divulgação de fake news e por crimes de ódio praticados por sistemas de inteligência artificial generativa / Vinicius de Fraga Bassotto. -- 2023.
102 f.
Orientador: Vanessa Chiari Gonçalves.

Trabalho de conclusão de curso (Graduação) --
Universidade Federal do Rio Grande do Sul, Faculdade de Direito, Curso de Ciências Jurídicas e Sociais, Porto Alegre, BR-RS, 2023.

1. Sistemas de inteligência artificial generativa. 2. Fake news. 3. Crimes de ódio. 4. Responsabilização penal. I. Gonçalves, Vanessa Chiari, orient. II. Título.

VINICIUS DE FRAGA BASSOTTO

**RESPONSABILIZAÇÃO PENAL PELA DIVULGAÇÃO DE *FAKE NEWS* E POR
CRIMES DE ÓDIO PRATICADOS POR SISTEMAS DE INTELIGÊNCIA ARTIFICIAL
GENERATIVA**

Trabalho de Conclusão de Curso
apresentado como requisito parcial para
obtenção do grau de Bacharel em Direito pela
Faculdade de Direito da Universidade Federal
do Rio Grande do Sul.

Porto Alegre, 15 de setembro de 2023.

BANCA EXAMINADORA:

Professora Doutora Ana Paula Motta Costa
Universidade Federal do Rio Grande do Sul

Professor Doutor Ângelo Roberto Ilha da Silva
Universidade Federal do Rio Grande do Sul

Professora Doutora Vanessa Chiari Gonçalves
Universidade Federal do Rio Grande do Sul

AGRADECIMENTOS

Este trabalho é composto pelo apoio de várias pessoas que passaram pela minha trajetória e que moldaram quem eu sou e meu jeito de enxergar o mundo.

Em primeiro lugar, dedico este trabalho aos meus pais, Vasco Antonio Bassotto e Vera Elisete de Fraga Bassotto. Mesmo em momentos de dificuldades, vocês jamais desistiram de fornecer estrutura e condições para contribuir com minha educação, principal legado que podem deixar.

Aos meus irmãos, Caroline de Fraga Bassotto e Vitor de Fraga Bassotto. À Caroline, irmã mais velha, agradeço por incentivar o desenvolvimento de meu senso crítico, que, inegavelmente, contribuiu para a escolha pelo caminho do Direito. Ao Vitor, irmão gêmeo, que participa da minha vida desde o “dia um”, agradeço por todo companheirismo e parceria, inclusive de anseios e preocupações.

Sem vocês, Vasco, Vera, Caroline e Vitor, nada do que fiz até hoje seria possível. Vocês foram – e são – minha maior fonte de amparo, possuidores de carinho incomensurável e, sem dúvida, as pessoas que mais desejam meu sucesso.

Ao meu namorado, Felipe Guedes Moreira Vieira, meu porto seguro, meu esteio, minha fonte de sorrisos e risos, que torna minha vida mais leve. Obrigado por compartilhar a vida comigo e me fazer respirar em momentos difíceis. Sem você, a escrita deste trabalho não teria sido possível.

Agradeço, ainda, aos meus amigos da Faculdade de Direito da UFRGS. Sou extremamente grato por, ao longo desses cinco anos – especialmente em meio a uma longa pandemia –, ter tido a possibilidade de conviver com pessoas tão especiais, sensíveis e parceiras. Compartilhar preocupações, medos e objetivos com vocês atenuou diversos desafios e contribuiu para construção de minha identidade.

Aos meus chefes e colegas do TJRS, do TRF4, da equipe criminal do Silveiro Advogados e do Aloísio Zimmer Advogados, que me introduziram ao mundo do Direito, me encorajaram a estar sempre atualizado e a desenvolver um trabalho de qualidade: muito obrigado pelos ensinamentos jurídicos, profissionais e, sobretudo, pelos conselhos de vida.

Por fim, porém não menos importante, dedico este trabalho a todos os professores que passaram pela minha trajetória, especialmente aqueles do Colégio Anchieta, do Colégio Monteiro Lobato e do Colégio Marista Champagnat: sem os ensinamentos de vocês, jamais teria alcançado o sonho de ingressar na concorrida Faculdade de Direito da UFRGS.

Ainda, dedico aos professores da Faculdade de Direito da UFRGS, que, mesmo em meio a tantos desafios, inerentes a uma Universidade pública – especialmente de recurso e de infraestrutura –, conseguem entregar um ensino, além de público e gratuito, de qualidade. Deixo um especial agradecimento à minha orientadora, Professora Doutora Vanessa Chiari Gonçalves, por sua disponibilidade, atenção e sensibilidade com seus orientandos.

Quem controla o passado, controla o futuro;
quem controla o presente, controla o passado.
(ORWELL, George. **1984**. p. 47)

RESUMO

O desenvolvimento de sistemas de inteligência artificial, em paralelo aos benefícios, tem propiciado o surgimento de uma grande variedade de riscos e de “efeitos colaterais negativos”. Nesse cenário, está a possibilidade de sua utilização como instrumento ou como agente independente para divulgar *fake news* e praticar discurso de ódio, fenômenos intrínsecos ao mundo polarizado e belicoso da era da internet e das redes sociais. Considerando a crescente automação e independência das tecnologias de inteligência artificial, passou-se ao questionamento quanto à possibilidade de os sistemas de inteligência artificial, notadamente, de inteligência artificial generativa, serem penalmente responsabilizados pela divulgação de *fake news* e pela prática de discurso de ódio. Inserido nesse contexto, o primeiro capítulo deste trabalho buscou analisar a problemática extrapenal ao serem tratadas as características, possibilidades de uso, benefícios, riscos e iniciativas de regulamentação dos sistemas de inteligência artificial, bem como verificou no que consiste o ChatGPT e os seus principais riscos. Em segundo lugar, ao se aproximar da problemática penal, procedeu-se à análise do tratamento conferido pelo direito penal brasileiro às *fake news*, no que consistem os crimes de ódio e de que forma os sistemas de inteligência artificial podem realizá-los. Por fim, sistematizamos as hipóteses de responsabilização penal em três modelos associados à maior ou à menor independência que possuem as tecnologias de inteligência artificial frente ao programador ou ao usuário. O primeiro modelo considera que um sujeito pode desenvolver ou utilizar das aptidões do sistema de inteligência artificial para, como instrumento, proferir discurso de ódio ou divulgar *fake news*, de modo que, nesta hipótese, responderia penalmente o próprio agente. O segundo modelo cogita a responsabilização a título de culpa, por negligência, na hipótese de o programador ou o desenvolvedor terem previsto, porém não terem impedido, a prática de ofensas de tais natureza. O terceiro modelo, concernente aos sistemas de inteligência artificial generativa, prevê a produção do resultado típico sem dependência com o programador ou o usuário, o que consiste no ponto sensível e nebuloso da matéria. Nessa hipótese, analisamos as noções de conduta, relação de causalidade e responsabilidade da pessoa jurídica. Ao final, o problema de pesquisa foi respondido em sentido contrário à possibilidade de responsabilização penal nos três casos: (I) do próprio sistema de inteligência artificial generativa, tendo em vista que não se trata de conduta à luz do direito penal; (II) do programador ou do usuário, considerando a inexistência de relação de causalidade entre conduta e resultado; e (III) da empresa desenvolvedora, haja vista a impossibilidade de responsabilização penal da pessoa jurídica nas hipóteses. Assim, abordamos alternativas à responsabilização penal, especialmente dentro dos campos da responsabilidade civil e administrativa, por ilícitos cometidos pelas tecnologias.

Palavras-chave: Sistemas de inteligência artificial generativa. *Fake news*. Crimes de ódio. Responsabilização penal.

ABSTRACT

The development of artificial intelligence entities, alongside with the benefits, has led to the emergence of risks and “negative side effects”. In this scenario, there is the possibility of using it as an instrument or as an independent agent to spread fake news and engage in hate speech, phenomena related to the polarized and bellicose world of the internet and social media era. Considering the increasing automation and independence of artificial intelligence technologies, questions have arisen regarding the possibility of artificial intelligence entities, notably generative artificial intelligence, being held criminally responsible for fake news dissemination and the practice of hate speech. Embedded in this context, the first chapter of this work sought to analyze the non-penal issue by addressing the characteristics, possibilities of use, benefits, risks, and regulatory initiatives of artificial intelligence entities, as well as examining what ChatGPT consists of and its main dangers. Secondly, when approaching the criminal problem, we analyzed the treatment given by Brazilian criminal law to fake news, searching for the definition of hate crimes and how artificial intelligence entities can carry them out. Finally, we systematize the hypotheses of criminal liability into three models associated with the greater or lesser independence of artificial intelligence technologies from the programmer or the user. The first model considers that a subject can develop or use the skills of the artificial intelligence entity to, as an instrument, utter hate speech or spread fake news, so in this hypothesis, the agent itself would be criminally liable. The second model considers liability for negligence if the programmer or developer had foreseen, but not prevented, the commission of offenses of such nature. The third model, concerning generative artificial intelligence entities, predicts the production of the typical result without dependence on the programmer or the user, which is the sensitive and nebulous point of the matter. In this hypothesis, we analyze the notions of conduct, causal link, and responsibility of the legal entity. In the end, the research problem was answered in the opposite direction to the possibility of criminal liability in the three cases: (I) the generative artificial intelligence entity itself, considering that it is not a question of conduct in the light of criminal law; (II) of the programmer or user, considering the non-existence of a causal relationship between conduct and result; and (III) the developer company, given the impossibility of holding the legal entity criminally liable in these cases. Thus, we will address alternatives to criminal liability for offenses committed by technologies, especially within the fields of civil liability and administrative law.

Keywords: Generative artificial intelligence entities. Fake news. Hate crimes. Criminal liability.

LISTA DE ABREVIATURAS E SIGLAS

IA	Inteligência Artificial
PL	Projeto de Lei
PLS	Projeto de Lei do Senado
CP	Código Penal
CF/88	Constituição Federal de 1988

SUMÁRIO

1 INTRODUÇÃO	11
2 CONTEXTUALIZAÇÃO DA PROBLEMÁTICA EXTRAPENAL: OS SISTEMAS DE INTELIGÊNCIA ARTIFICIAL	14
2.1 Definição e histórico	14
2.2 Possibilidades de uso	18
2.3 Benefícios e riscos da inteligência artificial	20
2.4 Regulamentação dos sistemas de inteligência artificial: a Lei de Inteligência Artificial da União Europeia e o Projeto de Lei n.º 2.388/2023	26
2.5 O ChatGPT	30
3 A DIVULGAÇÃO DE <i>FAKE NEWS</i> E A PRÁTICA DE CRIMES DE ÓDIO: UMA ANÁLISE À LUZ DO DIREITO PENAL	33
3.1 A divulgação de <i>fake news</i> e o tratamento pelo direito penal brasileiro	33
3.2 Os crimes de ódio em tempos de pós-modernidade	44
3.3 Os sistemas de inteligência artificial como divulgadores de <i>fake news</i> e praticantes de crimes de ódio	50
4 A (IM)POSSIBILIDADE DE RESPONSABILIZAÇÃO PENAL DE SISTEMAS DE INTELIGÊNCIA ARTIFICIAL GENERATIVA	54
4.1 Modelos de responsabilização de Hallevy	54
4.2 Aproximação da sistematização de Hallevy à dogmática penal brasileira	59
4.3 Particularidades da divulgação de <i>fake news</i> e da prática de crimes de ódio por sistemas de inteligência artificial	64
4.3.1 Noções introdutórias	65
4.3.2 A conduta	67
4.3.3 A relação de causalidade	71
4.3.4 A responsabilização penal da empresa desenvolvedora	75
4.3.5 Alternativas à responsabilização penal	81
5 CONCLUSÃO	87
REFERÊNCIAS	92

1 INTRODUÇÃO

Os sistemas de inteligência artificial vieram para ficar. Se, há não muito tempo, sistemas capazes de executar tarefas e ações autonomamente, sem intervenção humana, pareciam ser algo para um futuro distante, hoje podemos constatar que o futuro, na verdade, já é presente. O uso da inteligência artificial vem sendo rapidamente popularizado e difundido, com influência em um número cada vez maior de áreas, em ritmo e abrangência pouco vistos antes em outras tecnologias.

Conforme abordamos neste trabalho, a existência da inteligência artificial não é algo recente. Contudo, os últimos anos – quiçá, os últimos meses – levaram a uma expressiva ampliação na divulgação do conceito de “inteligência artificial”, mormente após o lançamento do ChatGPT, em novembro de 2022.

Até tempo atrás, “inteligência artificial” era um conceito pouco compreendido e discutido. Hoje, por outro lado, além de recorrente pauta na imprensa, está inserido em rodas de conversa. Nesse sentido, verifica-se o início de uma nova era, com a aproximação da inteligência artificial ao cotidiano e a sua inserção no vocabulário popular.

Com efeito, o expressivo avanço na utilização destes sistemas vem, sobremaneira, contribuindo com a otimização de tarefas e de atividades que, até pouco tempo, eram impensáveis ou consideradas de difícil execução. Ademais, o acesso ao universo virtual/digital transformou significativamente as relações humanas, visto que as tecnologias são capazes de diminuir distâncias e barreiras, acelerar a velocidade do compartilhamento de informações e permitir um número cada vez maior de conexões, com pessoas do mundo inteiro. A possibilidade ilimitada de conexões é, contudo, de modo paradoxal, diretamente proporcional ao quanto estamos, na prática, desconectados e distantes enquanto sociedade.

Notórias, ademais, as mudanças no que diz respeito ao acesso à informação. Se, há décadas, este era limitado e restrito a um determinado número de meios de comunicação e veículos jornalísticos, hoje, na atual “sociedade da informação”¹,

¹ “A sociedade da informação é a sociedade que está atualmente a constituir-se, na qual são amplamente utilizadas tecnologias de armazenamento e transmissão de dados e informação de baixo custo. Esta generalização da utilização da informação e dos dados é acompanhada por inovações organizacionais, comerciais, sociais e jurídicas que alterarão profundamente o modo de vida tanto no mundo do trabalho como na sociedade em geral. [...] Uma quantidade imensa de insumos informativos está à disposição nas redes (entre as quais ainda sobressai a Internet). Um grande número de agentes

encontram-se cada vez menos barreiras, e é possível o atingimento de um número cada vez maior de pessoas. Há não muitos anos, notícias chegavam exclusivamente por meio de jornais impressos, telejornais e revistas; hoje, é possível, quase que em tempo real, ter acesso ao que está acontecendo no mundo inteiro, literalmente na palma da mão, com o uso de *smartphones*.

Simultaneamente, o atual estágio tecnológico levou ao surgimento de um crescente número de desafios e riscos. Nesse cenário, dois dos principais “efeitos colaterais negativos”² que advieram da evolução tecnológica consistem na divulgação de *fake news* e na prática de crimes de ódio.

Os aparelhos eletrônicos e as redes sociais contribuíram para o desenvolvimento de um ambiente cada vez mais polarizado, conflituoso e agressivo, com a separação do “nós” *versus* “eles”. Nesse cenário, gênero, identidade de gênero, orientação sexual, cor da pele, etnia etc. passam a ser instrumentalizados para atacar e ofender quem pensa e age diferente do que se entende como certo e esperado.

Da mesma forma, divulgar inverdades e mentiras travestidas de notícias – nas chamadas *fake news* – passa a ser uma arma no belicoso conflito entre dois lados. Inequivocamente, tais efeitos colaterais negativos restaram amplificados com os sistemas de IA, em razão das suas características e potencialidades.

Com efeito, tem-se constatado o surgimento de sistemas de IA cada vez mais independentes, especialmente por meio de tecnologias de *machine learning*, chamados de modelos de inteligência artificial generativa. Nesses modelos, percebe-se a capacidade de adquirir aprendizado a partir de suas próprias experiências, executando ações sem terem sido direta ou especificamente programados para tanto. Na prática, tais modelos são aptos a “pensar” sozinhos, sem intervenção humana.

Por essa razão, a habilidade de “pensar” sozinho leva ao questionamento: na hipótese de os sistemas de inteligência artificial generativa praticarem um ilícito penal,

cognitivos humanos pode interligar-se em um mesmo processo de construção de conhecimentos.”. ASSMANN, Hugo. **A metamorfose do aprender na sociedade da informação. Ciência da Informação**, Brasília, v. 29, n. 2, p. 07-15, 2000. IBICT. Disponível em: <http://doi.org/10.1590/s0100-19652000000200002>. Acesso em: 23 jul. 2023. p. 8.

² NORVIG, Peter; RUSSEL, Stuart Jonathan. **Inteligência artificial: uma abordagem moderna**. 4. ed. Rio de Janeiro: LTC, 2022. Tradução de: Daniel Vieira e Flávio Soares Corrêa da Silva. *E-book*. Disponível em: <https://app.minhabiblioteca.com.br/#!/books/9788595159495>. Acesso em: 05 jul. 2023. p. 895.

eles poderão ser responsabilizados penalmente? Nesse sentido, com este trabalho, buscamos responder a esta problemática, aproximando-a, em especial, ao contexto da divulgação de *fake news* e da prática de crimes de ódio, tendo em vista a expressividade e o alcance de tais condutas no ambiente digital.

Desta forma, propomo-nos a responder ao seguinte questionamento: é possível responsabilizar penalmente sistemas de inteligência artificial, notadamente, modelos de inteligência artificial generativa, pela divulgação de *fake news* e pela prática de crimes de ódio por eles cometidos? Visando a responder a esta questão organizamos o presente trabalho em três partes.

Na primeira parte, tratamos da problemática extrapenal, abordando as principais características, possibilidades de uso e benefícios e riscos dos sistemas de IA. Da mesma forma, abordamos propostas de regulamentação dos sistemas, dando destaque à Lei de Inteligência Artificial da União Europeia e, no contexto brasileiro, ao Projeto de Lei n.º 2.388/2023. Por fim, realizamos exposição acerca da ferramenta “ChatGPT”, tendo em vista sua relevância para a aquisição de novos contornos no que tange às tecnologias de inteligência artificial generativa.

Na segunda parte, analisamos as questões das *fake news* e da prática de crimes de ódio à luz do direito penal. Nesse tópico, propomo-nos a examinar o tratamento que o direito penal brasileiro confere às *fake news*, bem como no que consistem os crimes de ódio no atual mundo digital. Da mesma forma, apontamos casos em que sistemas de IA divulgaram *fake news* e praticaram crimes de ódio, demonstrando, pois, a relevância prática da resolução do problema de pesquisa.

Na terceira e última parte, buscamos resolver o problema de pesquisa, visando a responder se é possível ou não a responsabilização penal de sistemas de inteligência artificial pela divulgação de *fake news* e pela prática de crimes de ódio. Para tanto, apresentamos a sistematização desenvolvida por Gabriel Hallevy, aproximando-a à dogmática penal brasileira. Da mesma forma, ao dar enfoque à análise da possibilidade de responsabilização penal no contexto da divulgação de *fake news* e da prática de crimes de ódio, tratamos acerca das noções de conduta, relação de causalidade e responsabilidade penal da pessoa jurídica, mais especificamente, da empresa desenvolvedora das tecnologias. Por fim, propusemos alternativas, fora do âmbito do direito penal, à responsabilização dos sistemas de IA por ilícitos praticados.

2 CONTEXTUALIZAÇÃO DA PROBLEMÁTICA EXTRAPENAL: OS SISTEMAS DE INTELIGÊNCIA ARTIFICIAL

2.1 Definição e histórico

O termo “inteligência artificial” foi cunhado e utilizado pela primeira vez por John McCarthy em conferência sobre o assunto por ele organizada, juntamente com Marvin Minsky, no Dartmouth College, em 1956 (“*Dartmouth Summer Research Project on Artificial Intelligence*”)³. Segundo Marvin Minsky, a inteligência artificial corresponde à “ciência de construir máquinas para que façam coisas que, se feitas por humanos, exigiriam inteligência”⁴.

A inteligência artificial pode também ser definida como o “estudo dos sistemas que agem de um modo que a um observador qualquer pareceria ser inteligente”⁵, envolvendo a utilização de métodos baseados no comportamento inteligente de humanos para resolução de problemas complexos⁶. São, em síntese, “máquinas com a capacidade de realizar funções como lógica, raciocínio, planejamento, aprendizagem e percepção”⁷.

Stuart Russell e Peter Norvig, na célebre obra “Inteligência Artificial”, sustentam que os sistemas de inteligência artificial correspondem a “máquinas que conseguem computar como agir de modo eficaz e seguro em uma grande variedade de novas situações.”⁸. A inteligência artificial, desta forma, se ocupa principalmente da ação

³ SILVA, Nilton Correia da. Parte I - Compreensão da inteligência artificial e dos seus pressupostos de controle e regulação. In: (Coord.) FRAZÃO, Ana; MULHOLLAND, Caitlin. **Inteligência artificial e direito: ética, regulação e responsabilidade**. 2. ed. São Paulo: Thomson Reuters Brasil, 2020. *E-book*. Disponível em: <https://proview.thomsonreuters.com/launchapp/title/rt/monografias/198921448>. Acesso em: 04 jul. 2023. p. RB-1.1.

⁴ “La Inteligencia Artificial es la ciencia de construir máquinas para que hagan cosas que, si las hicieran los humanos, requerirían inteligencia.”. ESPINOSA, Alejandra Morán. **Responsabilidad penal de la Inteligencia Artificial (IA)**. ¿La próxima frontera? Revista del Instituto de Ciencias Jurídicas de Puebla, México, Puebla, México, v. 15, n. 48, 2021. Disponível em: <https://doi.org/10.35487/rius.v15i48.2021.706>. Acesso em: 04 jul. 2023. p. 290 (tradução nossa).

⁵ COPPIN, Ben. **Inteligência artificial**. Rio de Janeiro: LTC, 2013. Tradução de: Jorge Duarte Pires. *E-book*. Disponível em: <https://app.minhabiblioteca.com.br/#/books/978-85-216-2936-8>. Acesso em: 04 jul. 2023. p. 4.

⁶ *Ibidem*.

⁷ SANTOS, Marcelo Henrique dos. **Introdução à inteligência artificial**. Londrina: Editora e Distribuidora Educacional S.A., 2021. *E-book*. Disponível em: <https://app.minhabiblioteca.com.br/#/books/9786559031245>. Acesso em: 04 jul. 2023. p. 6.

⁸ NORVIG, Peter; RUSSEL, Stuart Jonathan. **Inteligência artificial: uma abordagem moderna**. 4. ed. Rio de Janeiro: LTC, 2022. Tradução de: Daniel Vieira e Flávio Soares Corrêa da Silva. *E-book*. Disponível em: <https://app.minhabiblioteca.com.br/#/books/9788595159495>. Acesso em: 04 jul. 2023. p. 1.

racional, considerando a melhor ação possível a ser adotada em uma situação⁹.

Dora Kaufman¹⁰ ressalva, contudo, na linha da abordagem de Russel e Norvig, que as máquinas não são inteligentes no sentido conferido pelos seres humanos – isto é, ter a capacidade de agir para atingir objetivos próprios. Ao contrário, as máquinas não têm objetivos, pois são os seres humanos que os imputam, nada impedindo, porém, que as máquinas encontrem soluções melhores do que as encontradas por seres humanos. Por essa razão, propõe o abandono da ideia de “máquinas inteligentes” em favor de “máquinas benéficas”, ao se esperar que as ações das máquinas atinjam os objetivos dos seres humanos, melhorando a vida em sociedade. Segundo Stuart Russell e Peter Norvig¹¹:

Em resumo, a IA se concentra no estudo e na construção de agentes que fazem a coisa certa. Aquilo que é considerado a coisa certa é definido pelo objetivo que oferecemos ao agente. [...] Não queremos máquinas que sejam inteligentes no sentido de perseguir seus objetivos; queremos que elas busquem os nossos objetivos. Se não pudermos transferir esses objetivos perfeitamente para a máquina, então precisamos de uma nova formulação. [...] Em última análise, queremos agentes comprovadamente benéficos para os humanos.

Gabriel Hallevy¹² discorre acerca da evolução tecnológica ocorrida a partir de quando os computadores evoluem de máquinas “pensantes” – máquinas programadas a performar um determinado processo de pensamento computacional – para máquinas efetivamente pensantes, sem aspas, caracterizadas pelos sistemas de inteligência artificial. Para Hallevy, inteligência artificial é a capacidade de uma máquina imitar comportamento inteligente, isto é, é a simulação do comportamento humano e o processo cognitivo em um computador.

Historicamente, a origem do estudo dos sistemas da inteligência artificial remonta às décadas de 1940 e 1950. O primeiro trabalho reconhecido como IA foi realizado por Warren McCulloch e Walter Pitts, em 1943, ao proporem um modelo de

⁹ *Ibidem*, p. 31.

¹⁰ KAUFMAN, Dora. **Desmistificando a inteligência artificial**. Belo Horizonte: Autêntica, 2022. *E-book*. Disponível em: <https://app.minhabiblioteca.com.br/#/books/9786559281596>. Acesso em: 04 jul. 2023. p. 13.

¹¹ NORVIG, Peter; RUSSEL, Stuart Jonathan. **Inteligência artificial: uma abordagem moderna**. 4. ed. Rio de Janeiro: LTC, 2022. Tradução de: Daniel Vieira e Flávio Soares Corrêa da Silva. *E-book*. Disponível em: <https://app.minhabiblioteca.com.br/#/books/9788595159495>. Acesso em: 04 jul. 2023. p. 4.

¹² HALLEVY, Gabriel. **The Criminal Liability of Artificial Intelligence Entities - from Science Fiction to Legal Social Control**. Akron Intellectual Property Journal: v. 4, n. 2, 2010. Disponível em: <http://doi.org/10.2139/ssrn.1564096>. Acesso em: 04 jul. 2023. p. 2.

neurônios artificiais¹³.

No ano de 1950, Alan Turing – uma das maiores personalidades na história da inteligência artificial – publicou o trabalho “*Computing Machinery and Intelligence*”, no qual trata sobre o aprendizado de máquina, algoritmos genéticos e aprendizado por reforço. Ainda, no escrito, elaborou o “teste de Turing”, a fim de avaliar o sucesso ou não de uma tentativa de produzir um computador pensante: se uma pessoa interrogasse um computador e não conseguisse identificar se se tratava de um humano ou de um computador, o computador seria considerado, então, inteligente¹⁴.

Turing sustentava que seria mais fácil criar um sistema de inteligência artificial em nível humano desenvolvendo algoritmos de aprendizado e, em seguida, ensinando a máquina, em vez de programar sua inteligência manualmente. Mesmo mais de setenta anos depois, as ideias de Turing são atuais e ainda podem ser aplicadas aos sistemas de inteligência artificial hoje existentes. Adiante, a partir do final da década de 1980, passou a vigorar uma nova abordagem, mais científica, que incorpora a probabilidade e o aprendizado de máquina em vez de programação manual¹⁵.

A partir de 2001, iniciou-se o fenômeno do “*big data*”, que corresponde aos significativos avanços na computação e na criação da “*World Wide Web*”. Tais avanços viabilizaram a criação de conjuntos de dados enormes, com trilhões de palavras, bilhões de imagens e de horas de áudio e vídeo, bem como expressivas quantidades de dados. Como consequência direta está o desenvolvimento de algoritmos de aprendizado projetados usufruindo-se deste conjunto de dados¹⁶.

Mais recentemente, a datar de 2011, difundiram-se os métodos de aprendizagem profunda, consistentes no “aprendizado de máquina usando várias camadas de elementos de computação simples e configuráveis”¹⁷. Este modelo de aprendizado ocorreu mediante o reconhecimento de fala e, em sequência, no

¹³ NORVIG, Peter; RUSSEL, Stuart Jonathan. **Inteligência artificial: uma abordagem moderna**. 4. ed. Rio de Janeiro: LTC, 2022. Tradução de: Daniel Vieira e Flávio Soares Corrêa da Silva. *E-book*. Disponível em: <https://app.minhabiblioteca.com.br/#/books/9788595159495>. Acesso em: 05 jul. 2023. p. 16.

¹⁴ *Ibidem*, p. 2.

¹⁵ *Ibidem*, p. 22.

¹⁶ *Ibidem*, p. 23.

¹⁷ *Ibidem*, p. 24.

reconhecimento visual de objetos¹⁸.

Com efeito, durante anos existiu controvérsia acerca da essência do sistema de inteligência artificial, surgindo questões acerca da essência da humanidade – no sentido de se os seres humanos funcionam como máquinas pensantes – e dos sistemas de inteligência artificial – no sentido de se é possível existirem máquinas pensantes. Por essa razão, Gabriel Hallevy sistematizou cinco atributos esperáveis de um sistema inteligente: em primeiro lugar, a comunicação; em segundo lugar, o autoconhecimento; em terceiro lugar, o conhecimento externo; em quarto lugar, o comportamento orientado por objetivos; em quinto lugar, a criatividade¹⁹.

A propósito, um sistema de inteligência artificial é capaz de uma grande variedade de aplicações e possui capacidade de aprendizagem e ganho de experiência²⁰. Nesse sentido, sustenta-se que a inteligência artificial possui como características: autonomia, habilidade social, cooperação, proatividade e reatividade, os quais indicam a capacidade dos algoritmos em atuar sem – ou com pequena – intervenção humana²¹.

Atualmente, incrementa-se à complexidade dos sistemas de inteligência artificial por meio de modelos de *machine learning*, nos quais as máquinas têm aptidão de adquirir aprendizado a partir de suas próprias experiências²², isto é, “resume-se à capacidade de computadores executarem ações sem terem sido direta ou especificamente programados para tanto”²³. A crescente complexidade de problemas computacionais, a velocidade e o volume de dados levaram ao desenvolvimento de ferramentas baseadas em *machine learning*, ferramentas computacionais mais autônomas e independentes da intervenção humana para a aquisição de

¹⁸ *Ibidem*.

¹⁹ HALLEVY, Gabriel. **The Criminal Liability of Artificial Intelligence Entities - from Science Fiction to Legal Social Control**. Akron Intellectual Property Journal: v. 4, n. 2, 2010. Disponível em: <http://doi.org/10.2139/ssrn.1564096>. Acesso em: 04 jul. 2023. p. 5-6.

²⁰ *Ibidem*, p. 6.

²¹ TEPEDINO, Gustavo; SILVA, Rodrigo da Guia. **Desafios da inteligência artificial em matéria de responsabilidade civil**. Revista Brasileira de Direito Civil, Belo Horizonte, v. 21, n. 2, 2019. Disponível em: <https://rbdcivil.emnuvens.com.br/rbdc/article/view/465>. Acesso em: 04 jul. 2023. p. 63.

²² *Ibidem*.

²³ TEIDER, Lucas Hinckel; SANTOS, Gabriel Pivatto dos. **Responsabilidade penal da inteligência artificial (?): a problemática relacionada ao elemento da conduta na clássica estrutura analítica do delito**. In: (Coord.) GONZÁLEZ, Javier Garcia; LOZANO, Álvaro Alzina; RODRÍGUEZ, Gabriel Martín; GONÇALVES, Rubén Miranda; VEIGA, Fábio da Silva. *El derecho público y privado ante las nuevas tecnologías*. Madrid: Dykinson, 2020. Disponível em: <https://doi.org/10.2307/j.ctv153k4m>. Acesso em: 04 jul. 2023. p. 185-195.

conhecimento²⁴, também chamadas de sistemas de inteligência artificial generativa.

Segundo esquematiza Alejandra Morán Espinosa²⁵, há quatro modelos de inteligência artificial, quais sejam: (I) IA assistida, que colabora nas tarefas para realizá-las com rapidez; (II) IA automatizada, que realiza tarefas cotidianas de forma automática, geralmente de apoio administrativo; (III) IA aumentada, que facilita a tomada de decisões, aprendendo a interação realizada e os resultados obtidos; e (IV) IA autônoma, que possui a capacidade de tomada de decisões sem intervenção humana. Verifica-se, portanto, que os sistemas de inteligência artificial podem guardar maior ou menor independência ou necessidade de interação com um ser humano.

Nesse cenário, como resultado prático, tarefas que antes eram computacionalmente desafiadoras passaram a ser triviais e amplamente executadas diariamente. Dessa forma, evidencia-se uma constante expansão do uso da inteligência artificial para resolução de problemas reais, sendo cada vez maior os campos e contextos em que utilizados tais sistemas²⁶.

2.2 Possibilidades de uso

A ampliação da utilização de sistemas de IA vai ao encontro das demandas impostas por uma sociedade de consumo capitalista, que, constantemente, visa à otimização de processos, na medida em que são estas ferramentas capazes de sistematizar e automatizar tarefas intelectuais²⁷, que, se realizadas manualmente, demandariam maior investimento de tempo e dinheiro. A utilização da inteligência artificial encontra-se, nessa esteira, inserida dentro do contexto da globalização e sua revolução “técnica-científica-informacional”, conceito cunhado pelo geógrafo brasileiro

²⁴ FACELI, Katti; LORENA, Ana Carolina; GAMA, João; ALMEIDA, Tiago Agostinho de; CARVALHO, André C. P. L. F. de. **Inteligência artificial: uma abordagem de abordagem de aprendizado de máquina**. 2. ed. Rio de Janeiro: LTC, 2021. *E-book*. Disponível em: <https://app.minhabiblioteca.com.br/#/books/9788521637509>. Acesso em: 05 jul. 2023. p. 1.

²⁵ ESPINOSA, Alejandra Morán. **Responsabilidad penal de la Inteligencia Artificial (IA)**. ¿La próxima frontera? Revista del Instituto de Ciencias Jurídicas de Puebla, México, Puebla, México, v. 15, n. 48, 2021. Disponível em: <https://doi.org/10.35487/rius.v15i48.2021.706>. Acesso em: 04 jul. 2023. p. 293.

²⁶ *Ibidem*.

²⁷ PONTES, Marcos; ALVIM, Paulo Cesar Rezende de Carvalho; GONTIJO, José Gustavo Sampaio; AZAMBUJA, Eliana Cardoso Emediato de; VIDAL, Karina Domingues Bressan; OLIVEIRA, Cezar Luciano Cavalcanti de. **Inteligência artificial no contexto da estratégia brasileira de transformação digital**. In: (Coord.) VAINZOF, Rony; GUTIERREZ, Andriei. **Inteligência Artificial (IA)**. São Paulo: Revista dos Tribunais, 2021. *E-book*. Disponível em: <https://proview.thomsonreuters.com/launchapp/title/rt/monografias/270029414>. Acesso em: 05 jul. 2023. p. RB-1.2.

Milton Santos²⁸, que assevera:

A ciência, a tecnologia e a informação estão na base mesma de todas as formas de utilização e funcionamento do espaço, da mesma forma que participam da criação de novos processos vitais e da produção de novas espécies (animais e vegetais). É a cientificização e a tecnicização da paisagem. É, também, a informatização, ou, antes, a informacionalização do espaço. A informação tanto está presente nas coisas como é necessária à ação realizada sobre essas coisas. [...] Esses novos subespaços são, pois, mais ou menos capazes de rentabilizar uma produção.

Nesse sentido, verifica-se que os sistemas de inteligência artificial são capazes de beneficiar diversos segmentos, tais quais saúde e medicina, educação, transporte, segurança, indústria, agropecuária, informática, agricultura, judiciário, negócios, entretenimento, econômica, setor militar etc.²⁹. Com efeito, há muitos usos para os sistemas, tais como desempenhar atividades consideradas tediosas ou que requeiram que se dispenda tempo para realizá-las³⁰.

Conforme assinalam Fabiano Hartmann Peixoto e Roberta Zumblick Martins da Silva³¹, o conjunto de tarefas que os sistemas de inteligência artificial desempenham é muito mais amplo do que o de poucos anos atrás, incluindo:

[...] desde jogar jogos de tabuleiro a extrair os fatos de artigos de jornais, traduzir textos de uma língua para outra e o reconhecimento de discurso. É por causa da possibilidade de compreensão de texto – uma área denominada de processamento de linguagem natural [...] – que a IA se apresenta como um potencial para instrumentalizar soluções para problemas jurídicos.

A título exemplificativo, citam-se aplicações da inteligência artificial atualmente vislumbradas. Primeiramente, no campo dos transportes, destaca-se a existência de veículos autônomos e sem condutor, capazes de realizar transporte de passageiros por aplicativo, como da empresa Uber³². No campo da segurança digital, citam-se os

²⁸ SANTOS, Milton. **Técnica, espaço, tempo: globalização e meio técnico-científico informacional**. São Paulo: Hucitec, 1994. *E-book*. Disponível em: <http://geocrocetti.com/msantos/tecnica.pdf>. Acesso em: 05 jul. 2023. n. p.

²⁹ PONTES, Marcos; ALVIM, Paulo Cesar Rezende de Carvalho; GONTIJO, José Gustavo Sampaio; AZAMBUJA, Eliana Cardoso Emediato de; VIDAL, Karina Domingues Bressan; OLIVEIRA, Cezar Luciano Cavalcanti de. Inteligência artificial no contexto da estratégia brasileira de transformação digital. *In*: (Coord.) VAINZOF, Rony; GUTIERREZ, Andriei; RODRÍGUEZ, Gabriel Martín; GONÇALVES, Rubén Miranda; VEIGA, Fábio da Silva. **Inteligência Artificial (IA)**. São Paulo: Revista dos Tribunais, 2021. *E-book*. Disponível em: <https://proview.thomsonreuters.com/launchapp/title/rt/monografias/270029414>. Acesso em: 05 jul. 2023. p. RB-1.2.

³⁰ PEIXOTO, Fabiano Hartmann; SILVA, Roberta Zumblick Martins da. **Inteligência artificial e direito**. Curitiba: Alteridade Editora, 2019. p. 76.

³¹ *Ibidem*.

³² A propósito: UBER. **Autonomous mobility and delivery at Uber**. Disponível em: <https://www.uber.com/us/en/autonomous>. Acesso em: 04 jul. 2023.

algoritmos de classificação de mensagens de *spam*, os quais funcionam de maneira mais eficiente, tendo em vista a constante atualização de táticas de *spammers*, que dificulta abordagem estática programada³³.

No campo militar, verifica-se a existência de armas autônomas e robôs-soldados. Na área da saúde, sublinha-se a existência de robôs-médicos. No campo financeiro, mencionam-se os robôs com atuação autônoma em investimentos no mercado de ações. No judiciário, destacam-se sistemas autônomos de análise e de classificação processual, como o sistema “Victor”, no Supremo Tribunal Federal, apto à análise de recursos extraordinários recebidos pelo Tribunal, identificando vinculação com temas de repercussão geral³⁴.

Nos campos da comunicação e informacional, citam-se programas de comunicação de tradução automática e a utilização de “*chatbots*” inteligentes, como o ChatGPT – aplicação frequentemente associada à inteligência artificial –, assim como, no entretenimento, programas de imitação de imagem e de fala de outros indivíduos. Por fim, no campo das redes sociais, insta mencionar a existência de algoritmos de recomendação e sugestão de conteúdo.

2.3 Benefícios e riscos da inteligência artificial

A partir do crescente uso da inteligência artificial, uma série de benefícios a diversos segmentos econômicos e sociais surgiram, em uma abrangência pouco vislumbrada em demais tecnologias, tendo em vista o aperfeiçoamento e a agilização de diversas atividades humanas. Entretanto, em paralelo ao advento de benefícios, passam-se a verificar riscos e “efeitos colaterais negativos”³⁵ aos usuários na utilização de sistemas de inteligência artificial.

No que tange aos benefícios, é possível verificá-los em três dimensões³⁶: a

³³ NORVIG, Peter; RUSSEL, Stuart Jonathan. **Inteligência Artificial**. Rio de Janeiro: LTC, 2013. Tradução de Regina Célia Simille de Macedo. *E-book*. Disponível em: <https://app.minhabiblioteca.com.br/#/books/9788595156104>. Acesso em: 05 jul. 2023. p. 25.

³⁴ TEPEDINO, Gustavo; SILVA, Rodrigo da Guia. **Desafios da inteligência artificial em matéria de responsabilidade civil**. Revista Brasileira de Direito Civil, Belo Horizonte, v. 21, n. 2, 2019. Disponível em: <https://rbdcivil.emnuvens.com.br/rbdc/article/view/465>. Acesso em: 05 jul. 2023. p. 66.

³⁵ NORVIG, Peter; RUSSEL, Stuart Jonathan. **Inteligência artificial: uma abordagem moderna**. 4. ed. Rio de Janeiro: LTC, 2022. Tradução de: Daniel Vieira e Flávio Soares Corrêa da Silva. *E-book*. Disponível em: <https://app.minhabiblioteca.com.br/#/books/9788595159495>. Acesso em: 05 jul. 2023. p. 895.

³⁶ STEIBEL, Fabro; VICENTE, Victor Freitas; JESUS, Diego Santos Viera de. Possibilidades e potenciais da utilização da Inteligência Artificial. In: (Coord.) FRAZÃO, Ana; MULHOLLAND, Caitlin. **Inteligência artificial e direito: ética, regulação e responsabilidade**. 2. ed. São Paulo:

primeira, consistente na organização de dados; a segunda, refere-se ao auxílio à tomada de decisões; a terceira, por sua vez, diz respeito à automação de decisões. Com relação à organização, a título de exemplo, empresas podem organizar e gerir portfólios de clientes, enquanto governos podem sistematizar dados de cidadãos para elaboração de políticas públicas.

Por sua vez, o auxílio à tomada de decisões tem íntima relação com o processamento e a oferta de informações fundamentais visando ao conhecimento, de modo que o desenvolvimento de sistemas de inteligência artificial pode ser útil a gestores e empresas no processo de tomada de decisões, inclusive no desenvolvimento de produtos, estratégias de marketing e vendas. A terceira dimensão, por sua vez, assume destaque ainda maior: outrora, a tecnologia tradicional apenas auxiliava no processo de informações e comunicações, sendo a decisão, ao final, tomada por um ser humano. Com os sistemas de inteligência artificial, os próprios computadores são capazes de tomar decisões apropriadas – nos campos empresarial, de vendas, de marketing etc.

Com efeito, consoante o Relatório de Regulação da Inteligência Artificial, elaborado pela Escola Nacional de Administração Pública, na esteira da Proposta de Regulamento da Inteligência Artificial do Parlamento Europeu³⁷:

A IA é capaz de oferecer um vasto conjunto de benefícios econômicos e sociais a todo o leque de indústrias e atividades sociais ao melhorar as previsões, otimizar operações, a repartição de recursos e personalizar soluções digitais disponibilizadas às pessoas e às organizações. A utilização da IA também pode conferir importantes vantagens competitivas às empresas e contribuir para progressos sociais e ambientais. Ela pode ser utilizada, por exemplo, nos cuidados de saúde, na agricultura, na educação, no transporte e logística, na segurança, na eficiência energética, dentre muitas outras possibilidades³⁸.

Em síntese, conforme sustentam Russel e Norvig³⁹, considerando que toda a

Thomson Reuters Brasil, 2020. *E-book*. Disponível em: <https://proview.thomsonreuters.com/launchapp/title/rt/monografias/198921448>. Acesso em: 04 jul. 2023. p. RB-2.3.

³⁷ COMISSÃO EUROPEIA. **Regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial (regulamento inteligência artificial) e altera determinados atos legislativos da União**. Bruxelas, 2021. Disponível em: <https://eur-lex.europa.eu/legal-content/PT/TXT/?uri=CELEX:52021PC0206>. Acesso em: 04 jul. 2023.

³⁸ MELO, Ana Karolina Acris; SOUZA, Géssica Cardoso; VASCO, Amanda Corrêa; REIS, Bruno Salomon. **Regulação da inteligência artificial: benchmarking de países selecionados**. [S.L.]: Escola Nacional de Administração Pública, 2022. Disponível em: <http://repositorio.enap.gov.br/handle/1/7419>. Acesso em: 04 jul. 2023. p. 5.

³⁹ NORVIG, Peter; RUSSEL. Stuart Jonathan. **Inteligência artificial: uma abordagem moderna**. 4. ed. Rio de Janeiro: LTC, 2022. Tradução de: Daniel Vieira e Flávio Soares Corrêa da Silva. *E-book*.

civilização é resultado da inteligência humana, o acesso a uma inteligência mecânica maior permitirá a elevação do teto de ambições. A inteligência artificial, destarte, possui potencial de libertação da humanidade de trabalhos braçal e repetitivo, contribuindo para o aumento da produção de bens e serviços.

Os benefícios, contudo, coexistem com uma série de riscos e pontos de preocupação a serem observados. Tais riscos, até certo tempo, inexistiam, de modo que passaram a ser vislumbrados a partir do contexto da Quarta Revolução Industrial, a qual, segundo Klaus Schwab⁴⁰, é “caracterizada por uma internet mais ubíqua e móvel, por sensores menores e mais poderosos que se tornaram mais baratos e pela inteligência artificial e aprendizagem automática”, cujas tecnologias, “mais sofisticadas e integradas, [estão] transformando a sociedade e a economia global”.

Thomas Dietterich e Eric Horvitz⁴¹ sustentam que há cinco classes de riscos abrangendo o uso de sistemas de inteligência artificial: a primeira classe consiste nas falhas (“*bugs*”); a segunda classe refere-se aos riscos de segurança (“*cybersecurity*”), considerando a vulnerabilidade de ataques cibernéticos destes sistemas. A terceira, traduzida como “aprendiz de feiticeiro” (“*sorcerer’s apprentice*”), refere-se ao raciocínio que deve ser feito pelos sistemas de IA, o qual deve “analisar e compreender se o comportamento que um ser humano está solicitando pode ser julgado como ‘normal’ ou ‘razoável’ pela maioria das pessoas”⁴². A quarta classe denomina-se “autonomia compartilhada” (“*shared autonomy*”), que consiste nos riscos envolvendo a construção de um sistema colaborativo, com desafios na determinação nos objetivos do sistema. Por fim, como quinta categoria, existem os impactos socioeconômicos (“*influences on socioeconomics and the distribution of wealth*”).

Nesse sentido, faz-se oportuna a referência a exemplos de riscos:

Disponível em: <https://app.minhabiblioteca.com.br/#!/books/9788595159495>. Acesso em: 05 jul. 2023. p. 28.

⁴⁰ SCHWAB, Klaus. **A Quarta Revolução Industrial**. São Paulo: Edipro, 2019. *E-book*. Disponível em: <https://books.google.com.br/books?id=XZSWDwAAQBAJ>. Acesso em: 05 jul. 2023. n. p.

⁴¹ SICHMAN, Jaime Simão. **Inteligência Artificial e sociedade: avanços e riscos**. Estudos Avançados, São Paulo, v. 35, 2021. Disponível em: <https://doi.org/10.1590/s0103-4014.2021.35101.004>. Acesso em: 05 jul. 2023. p. 37–50.; DIETTERICH, Thomas G.; HORVITZ, Eric J. **Rise of concerns about AI**. Communications Of The Acm, v. 58, n. 10, 2015. Disponível em: <https://doi.org/10.1145/2770869>. Acesso em: 05 jul. 2023. p. 38-40.

⁴² “An AI system must analyze and understand whether the behavior that a human is requesting is likely to be judged as ‘normal’ or ‘reasonable’ by most people.”. DIETTERICH, Thomas G.; HORVITZ, Eric J. **Rise of concerns about AI**. Communications Of The Acm, v. 58, n. 10, 2015. Disponível em: <https://doi.org/10.1145/2770869>. Acesso em: 05 jul. 2023. p. 38-40.

I. Tomada de decisões tendenciosas⁴³: Algoritmos de aprendizado de máquina podem ser empregados em tarefas decisórias, como avaliação de pedidos de empréstimo, e tomar decisões discriminatórias com base na raça, sexo etc. No entanto, há quem sustente que a existência de vieses e preconceito está associada à conduta humana, sendo que, “no atual estágio, em que as máquinas dependem de supervisão humana, cabe aos profissionais da computação inserir nas máquinas os parâmetros corretos, ou seja, a responsabilidade sobre o processo⁴⁴.”. Atualmente, fala-se, inclusive, na existência de uma forma de manifestação do racismo conhecida como “racismo algorítmico”, conceito definido por Tarcízio Silva como⁴⁵:

O modo pelo qual a disposição de tecnologias e imaginários sociotécnicos em um mundo moldado pela supremacia branca realiza a ordenação algorítmica racializada de classificação social, recursos e violência em detrimento de grupos minorizados. Tal ordenação pode ser vista como uma camada adicional do racismo estrutural, que vai além e molda o futuro e horizontes de relações de poder adicionando mais opacidade sobre exploração e opressão global racializada desde o projeto colonial do século XVI.

II. Impactos nas relações de emprego⁴⁶: Os sistemas de inteligência artificial podem eliminar determinadas modalidades de trabalho, que passarão a ser realizadas por máquinas. Sustenta-se que as máquinas podem viabilizar economicamente algumas atividades que seriam inviáveis, resultando, como consequência, aumento de riqueza, o que poderá ter como efeito a transferência da riqueza do trabalho para o capital e o aumento de desigualdades. Felipe Barcarollo⁴⁷, de modo mais otimista, destaca que, em razão da complexidade da comunicação, existem tarefas que ainda só o ser humano poderá executar. A própria criação de algoritmos de inteligência artificial necessita de intervenção humana, já que não podem as máquinas, por si próprias, se autoprogramarem.

⁴³ NORVIG, Peter; RUSSEL. Stuart Jonathan. **Inteligência artificial: uma abordagem moderna**. 4. ed. Rio de Janeiro: LTC, 2022. Tradução de: Daniel Vieira e Flávio Soares Corrêa da Silva. *E-book*. Disponível em: <https://app.minhabiblioteca.com.br/#/books/9788595159495>. Acesso em: 05 jul. 2023. p. 29.

⁴⁴ KAUFMAN, Dora. **A inteligência artificial irá suplantará a inteligência humana?** Barueri: Estação das Letras e Cores, 2019. p. 42.

⁴⁵ SILVA, Tarcízio. **Racismo algorítmico: inteligência artificial e discriminação nas redes digitais**. São Paulo: Edições Sesc, 2022. *E-book*. Disponível em: <https://books.google.com.br/books?id=BCZcEAAAQBA>. Acesso em: 05 jul. 2023. n. p.

⁴⁶ NORVIG, Peter; RUSSEL. Stuart Jonathan. **Inteligência artificial: uma abordagem moderna**. 4. ed. Rio de Janeiro: LTC, 2022. Tradução de: Daniel Vieira e Flávio Soares Corrêa da Silva. *E-book*. Disponível em: <https://app.minhabiblioteca.com.br/#/books/9788595159495>. Acesso em: 05 jul. 2023. p. 29.

⁴⁷ BARCAROLLO, Felipe. **Inteligência artificial: aspectos ético-jurídicos**. São Paulo: Almedina, 2021. *E-book*. Disponível em: <https://app.minhabiblioteca.com.br/#/books/9786556272801>. Acesso em: 05 jul. 2023. p. 238.

III. Preocupações com a privacidade, na medida em que as tecnologias de IA usualmente procedem à coleta e à análise de grandes quantidades de dados pessoais.

IV. Riscos de segurança⁴⁸: Os sistemas de IA, cada vez mais avançados, podem ser utilizados como instrumentos para o desenvolvimento de ataques cibernéticos por hackers.

V. Contribuição para desinformação e manipulação de informações: Sistemas de inteligência artificial podem contribuir para a disseminação de informações falsas e a manipulação da opinião pública, através de mecanismos como os *deepfakes*. Michel K. Spencer assim descreve estes mecanismos⁴⁹:

Deepfakes são, essencialmente, identidades falsas criadas com o *Deep Learning* [aprendizagem profunda, por meio de uso maciço de dados], por meio de uma técnica de síntese de imagem humana baseada na inteligência artificial. É usada para combinar e sobrepor imagens e vídeos preexistentes e transformá-los em imagens ou vídeos “originais” [...] Essa combinação de vídeos existentes e “originais” resulta em vídeos falsos, que mostram uma ou algumas pessoas realizando ações ou fazendo coisas que nunca aconteceram na realidade.

Tais mecanismos de *deepfakes* são aptos a criar situações embaraçosas e disseminar conteúdos e informações falsas, sendo que:

[...] o crescimento dos *deepfakes* apresentou uma nova concepção à rede social virtual, pois essa nova inteligência artificial trouxe ao cotidiano mundial a desinformação, que tende a se aprimorar, colocando em risco a relação de confiança entre o ser humano e a internet⁵⁰.

Acrescenta-se, por fim, outro efeito colateral negativo decorrente do desenvolvimento dos sistemas de IA, consistente no surgimento de uma “nova criminalidade”. A “nova criminalidade” surgiu a partir de fatores da globalização da economia, bem como pela evolução tecnológica, comercial, legal e cultural, sendo que “práticas ilícitas causadoras da vitimização difusa buscam apoio no emprego da moderna tecnologia via internet e telefonia celular”⁵¹.

⁴⁸ NORVIG, Peter; RUSSEL, Stuart Jonathan. **Inteligência artificial: uma abordagem moderna**. 4. ed. Rio de Janeiro: LTC, 2022. Tradução de: Daniel Vieira e Flávio Soares Corrêa da Silva. *E-book*. Disponível em: <https://app.minhabiblioteca.com.br/#/books/9788595159495>. Acesso em: 05 jul. 2023. p. 29.

⁴⁹ SPENCER, Michael K.. **Deep Fake, a mais recente ameaça distópica**. 2019. Tradução de: Gabriela Leite. Disponível em: <https://ihu.unisinos.br/78-noticias/589686-deep-fake-a-mais-recente-ameaca-distopica>. Acesso em: 05 jul. 2023.

⁵⁰ ROBLES-LESSA, Moyana Mariano; CABRAL, Hildeliza Lacerca Tinoco Boechat; SILVESTRE, Gilberto Fachetti. **Deepfake: a inteligência artificial e o algoritmo causando riscos à sociedade no ciberespaço**. *Derecho y Cambio Social*, Lima-Peru, n. 61, Jul./Set., 2020. p. 480.

⁵¹ DUARTE, Maria Carolina de Almeida. **Globalização e a Nova Criminalidade**. *Revista Territórios e Fronteiras*, Cuiabá, v. 2, n. 1, 2011. *Revista Territórios e Fronteiras*. Disponível em:

Com efeito, os sistemas de inteligência artificial podem ser empregados como ferramentas – ou mesmo agentes – para prática de crimes, utilizando-se de suas capacidades para facilitar ações contra alvos do mundo real; nesses casos, ainda que, muitas vezes, os métodos sejam novos, podem possuir natureza tradicional – isto é, são crimes de roubo, extorsão, intimidação etc. Além disso, sistemas de inteligência artificial podem ser alvos de atividade criminal, bem como podem simplesmente fornecer contexto para a prática de crimes⁵².

As atividades criminais podem variar grandemente, tendo como alvos indivíduos ou instituições, negócios ou consumidores, governos etc., que podem ser motivados por diferentes fatores. Não bastasse, o ambiente digital contribui para a prática de crimes, os quais, no mundo digital, diferente dos crimes tradicionais:

[...] costumam ser altamente replicáveis: uma vez desenvolvidas, as técnicas podem ser compartilhadas, repetidas e até vendidas, permitindo o potencial para comercialização de técnicas criminosas ou fornecimento de ‘crime como serviço’. Isso pode levar a uma redução das barreiras tecnológicas, pois os criminosos podem externalizar os aspectos mais desafiadores de seus crimes baseados em IA⁵³.

Matthew Caldwell, Jerome Andrews, Thomas Tanay e Lewis Griffin, do University College London, a propósito, elaboraram estudo identificando quais crimes envolvendo sistemas de inteligência artificial possuem maior risco e devem ser objeto de maior preocupação⁵⁴. Na categoria de risco alto, foram incluídos: imitações audiovisuais, veículos sem condutores utilizados como armas, *phishing*⁵⁵ sob medida, perturbação de sistemas controlados por inteligência artificial, chantagens em larga

<https://periodicoscientificos.ufmt.br/territoriosefronteiras/index.php/v03n02/article/view/32/31>. Acesso em: 05 jul. 2023. p. 90.

⁵² CALDWELL, Matthew; ANDREWS, Jerone. T. A.; TANAY, Thomas; GRIFFIN, Lewis. D.. **AI-enabled future crime**. Crime Science, [S.L.], v. 9, n. 1, 2020. Springer Science and Business Media LLC. Disponível em: <https://doi.org/10.1186/s40163-020-00123-8>. Acesso em: 05 jul. 2023. p. 5.

⁵³ “Moreover, unlike many traditional crimes, crimes in the digital realm are often highly replicable: once developed, techniques can be shared, repeated, even sold, allowing the potential for marketisation of criminal techniques or provision of ‘crime as a service’. This may lead to a lowering of technological barriers as criminals are able to outsource the more challenging aspects of their AI-based crime.” *Ibidem* (tradução nossa).

⁵⁴ *Ibidem*.

⁵⁵ *Phishing* consiste em um ataque de engenharia social que visa à coleta de informações de segurança ou instalação de malware por meio de uma mensagem que pretende ser de uma parte confiável, como banco do usuário. O invasor explora a confiança existente para persuadir o usuário a realizar ações das quais ele poderia desconfiar, como revelar senhas ou clicar em links duvidosos. (“*Phishing* is a social engineering attack that aims to collect secure information or install malware via a digital message purporting to be from a trusted party such as the user’s bank. The attacker exploits the existing trust to persuade the user to perform actions they might otherwise be wary of, such as revealing passwords or clicking on dubious links.”). *Ibidem*, p. 7 (tradução nossa).

escala, *fake news* produzidas por IA⁵⁶.

Na categoria de risco médio, inseriram-se: robôs militares, venda de serviços fraudulentos, manipulação de dados, ataques em sistemas de aprendizagem, ataques de drones controlados de maneira autônoma, fraude de sistemas de reconhecimento facial, manipulação dos mercados financeiro e de ações. Por sua vez, na categoria de risco baixo, foram incluídos: exploração de vieses tendenciosos, *bots* ladrões, realização de avaliações falsas por sistemas de inteligência artificial, *stalking* e falsificação de conteúdo⁵⁷.

2.4 Regulamentação dos sistemas de inteligência artificial: a Lei de Inteligência Artificial da União Europeia e o Projeto de Lei n.º 2.388/2023

À medida que surgem e são popularizadas novas tecnologias, tem-se verificado um crescente aumento da preocupação dos Estados em regulamentar a inteligência artificial. Pretende-se com a regulamentação assegurar o uso de maneira segura, responsável, transparente, auditável, em respeito à privacidade de dados e aos direitos humanos, evitando a utilização para fins indevidos e ilícitos, outrora abordados neste trabalho.

Nesse cenário, sobleva destacar a elaboração de leis nos contextos europeu e brasileiro, precursores em propor uma efetiva regulamentação do uso da inteligência artificial. Em primeiro lugar, destaca-se a aprovação pelo Parlamento Europeu, no dia 14 de junho de 2023, de projeto de lei que regulamenta o uso de sistemas de IA na União Europeia, primeiro do gênero do mundo.

Consoante sustenta o Parlamento Europeu⁵⁸:

A prioridade do Parlamento é garantir que os sistemas de IA utilizados na UE sejam seguros, transparentes, rastreáveis, não discriminatórios e respeitadores do ambiente [de modo que] devem ser supervisionados por pessoas, em vez de serem automatizados, para evitar resultados prejudiciais. O Parlamento quer também estabelecer uma definição uniforme e neutra em termos tecnológicos para a IA, de modo a ser aplicada em futuros sistemas de IA.

⁵⁶ *Ibidem*, p. 6-9.

⁵⁷ *Ibidem*, p. 9-11.

⁵⁸ PARLAMENTO EUROPEU. **Lei da UE sobre IA: primeira regulamentação de inteligência artificial.** 2023. Disponível em: <https://www.europarl.europa.eu/news/pt/headlines/society/20230601STO93804/lei-da-ue-sobre-ia-primeira-regulamentacao-de-inteligencia-artificial>. Acesso em: 04 jul. 2023.

Assevera ainda:

As regras visam promover a adoção de uma IA centrada no ser humano e fiável, bem como proteger a saúde, a segurança, os direitos fundamentais e a democracia dos seus efeitos nocivos [, bem como] assegurarão que a IA desenvolvida e utilizada na Europa respeita plenamente os direitos e valores da UE, incluindo a supervisão humana, a segurança, a privacidade, a transparência, a não discriminação e o bem-estar social e ambiental⁵⁹.

Dessa forma, as regras elaboradas estabeleceram obrigações aos fornecedores e usuários de acordo com o nível de risco da inteligência artificial. Assim, os sistemas de inteligência artificial com um nível de risco inaceitável para a segurança das pessoas serão proibidos, na medida em que considerados uma ameaça para as pessoas, a exemplo de sistemas cujo propósito é de classificar as pessoas com base em seu comportamento, classe socioeconômica ou características pessoais⁶⁰. Por sua vez, os sistemas de risco elevado – como sistemas utilizados para influenciar eleitores, resultados de eleições e sistemas de recomendação das plataformas de redes sociais – devem ser avaliados antes de serem colocados no mercado e durante todo o seu ciclo de vida⁶¹.

Adiante, denominou-se a chamada “inteligência artificial generativa”, a exemplo do ChatGPT. Ferramentas inseridas nesta denominação têm o dever de cumprir com os requisitos de transparência, sinalizando que os conteúdos foram gerados por inteligência artificial, diferenciando técnicas de manipulação de imagens das imagens reais, bem como devem assegurar a geração de conteúdos ilegais.

As atividades de risco limitado⁶², por sua vez, devem cumprir requisitos mínimos de transparência, ao permitir que os usuários tomem decisões informadas, sendo que, após a interação, o usuário pode decidir se quer continuar utilizando ou não. Ainda, devem ser alertados para o fato de estarem interagindo com a IA.

Por seu turno, no Brasil, na data de 03 de maio de 2023, foi apresentado o Projeto de Lei n.º 2338, de 2023, que “dispõe sobre o uso da Inteligência Artificial”.

⁵⁹ PARLAMENTO EUROPEU. **Parlamento negocia primeiras regras para inteligência artificial mais segura.** 2023. Disponível em: <https://www.europarl.europa.eu/news/pt/press-room/20230609IPR96212/parlamento-negoceia-primeiras-regras-para-inteligencia-artificial-mais-segura>. Acesso em: 04 jul. 2023.

⁶⁰ PARLAMENTO EUROPEU. **Lei da UE sobre IA: primeira regulamentação de inteligência artificial.** 2023. Disponível em: <https://www.europarl.europa.eu/news/pt/headlines/society/20230601STO93804/lei-da-ue-sobre-ia-primeira-regulamentacao-de-inteligencia-artificial>. Acesso em: 04 jul. 2023.

⁶¹ *Ibidem*.

⁶² *Ibidem*.

Referido PL – conhecido como “Marco Legal da Inteligência Artificial no Brasil” – foi elaborado pela comissão especial de juristas presidida pelo ministro do Superior Tribunal de Justiça, Ricardo Villas Bôas Cueva, e apresentado pelo Senador Rodrigo Pacheco (PSD-MG)⁶³.

Ao dispor sobre o uso da inteligência artificial, o Projeto de Lei, em seu artigo 1º, refere que⁶⁴:

Art. 1º Esta Lei estabelece normas gerais de caráter nacional para o desenvolvimento, implementação e uso responsável de sistemas de inteligência artificial (IA) no Brasil, com o objetivo de proteger os direitos fundamentais e garantir a implementação de sistemas seguros e confiáveis, em benefício da pessoa humana, do regime democrático e do desenvolvimento científico e tecnológico.

O artigo 2º do PL, por sua vez, prevê em quais fundamentos o desenvolvimento, a implementação e o uso dos sistemas de inteligência artificial no Brasil estarão pautados, dentre eles: a centralidade da pessoa humana (inciso I); o respeito aos direitos humanos e aos valores democráticos (inciso II); a igualdade, a não discriminação, a pluralidade e o respeito aos direitos trabalhistas (inciso V); a privacidade, a proteção de dados e autodeterminação informativa (inciso VIII); e o acesso à informação e à educação, e a conscientização sobre os sistemas de inteligência artificial e suas aplicações (inciso X); etc.

Adiante, o artigo 3º prevê que o desenvolvimento, a implementação e o uso dos sistemas de inteligência artificial, além da boa-fé, deverão observar princípios como: crescimento inclusivo, desenvolvimento sustentável e bem-estar (inciso I); autodeterminação e liberdade de decisão e de escolha (inciso II); participação humana no ciclo da inteligência artificial e supervisão humana efetiva (inciso III); não discriminação (inciso IV); justiça, equidade e inclusão (inciso V); transparência, explicabilidade, inteligibilidade e auditabilidade (inciso V); rastreabilidade das decisões durante o ciclo de vida de sistemas de inteligência artificial como meio de prestação de contas e atribuição de responsabilidades a uma pessoa natural ou jurídica (inciso IX); etc. Oportuno salientar a previsão do inciso IX, considerando a proximidade com

⁶³ SENADO FEDERAL. **Uso de inteligência artificial pelo poder público será sujeito a regulamentação.** 2023. Disponível em: <https://www12.senado.leg.br/noticias/materias/2023/05/12/uso-de-inteligencia-artificial-pelo-poder-publico-sera-sujeito-a-regulamentacao>. Acesso em: 04 jul. 2023.

⁶⁴ SENADO FEDERAL. Dispõe sobre o uso da Inteligência Artificial. **Projeto de Lei Nº 2338, de 2023.** Disponível em: <https://www25.senado.leg.br/web/atividade/materias/-/materia/157233>. Acesso em: 04 jul. 2023.

o tema do presente trabalho.

O inciso I do artigo 4º, em sentido análogo às definições expostas alhures, traz completa e detalhada definição acerca do que consiste no sistema de inteligência artificial, definido como:

Sistema de inteligência artificial: sistema computacional, com graus diferentes de autonomia, desenhado para inferir como atingir um dado conjunto de objetivos, utilizando abordagens baseadas em aprendizagem de máquina e/ou lógica e representação do conhecimento, por meio de dados de entrada provenientes de máquinas ou humanos, com o objetivo de produzir previsões, recomendações ou decisões que possam influenciar o ambiente virtual ou real.

Os artigos 5º a 12 tratam acerca dos direitos dos usuários de sistemas inteligência artificial, arrolados no artigo 5º:

- I** – direito à informação prévia quanto às suas interações com sistemas de inteligência artificial;
- II** – direito à explicação sobre a decisão, recomendação ou previsão tomada por sistemas de inteligência artificial;
- III** – direito de contestar decisões ou previsões de sistemas de inteligência artificial que produzam efeitos jurídicos ou que impactem de maneira significativa os interesses do afetado;
- IV** – direito à determinação e à participação humana em decisões de sistemas de inteligência artificial, levando-se em conta o contexto e o estado da arte do desenvolvimento tecnológico;
- V** – direito à não-discriminação e à correção de vieses discriminatórios diretos, indiretos, ilegais ou abusivos; e
- VI** – direito à privacidade e à proteção de dados pessoais, nos termos da legislação pertinente.

Os artigos 13 a 18, ao encontro da legislação europeia, dispõem acerca da categorização dos riscos, a qual deve ser realizada previamente à colocação no mercado ou utilização em serviço, descrevendo o que consistem em sistemas de risco excessivo – cuja implementação deve ser vedada – e sistemas de alto risco. Os artigos 19 a 26, a seu turno, tratam da estrutura de governança que deverá existir para garantir a segurança dos sistemas e o atendimento dos direitos de pessoas afetadas.

Ainda, de acordo com os artigos 32 a 35, a supervisão e a fiscalização dos sistemas de inteligência artificial será responsabilidade de uma autoridade competente designada pelo Poder Executivo, a quem caberá, dentre outras funções, zelar pela proteção aos direitos fundamentais (inciso I); estimular a adoção de boas práticas no setor (inciso IV); expedir normas para regulamentação da Lei (inciso VI); e aplicar sanções administrativas no caso de descumprimento da lei (inciso IX).

Por fim, destacam-se dois dispositivos de sobeja relevância a este trabalho –

cuja análise, posteriormente, será aprofundada –, quais sejam, os artigos 27 e 36; o primeiro trata acerca de hipóteses de responsabilização civil do “fornecedor ou operador de sistema de inteligência artificial que cause dano patrimonial, moral, individual ou coletivo, [o qual será] obrigado a repará-lo integralmente, independentemente do grau de autonomia do sistema”, enquanto o segundo trata das sanções administrativas que podem ser aplicadas aos agentes de inteligência artificial em razão das infrações cometidas às normas previstas na Lei.

A propósito, salutar a relevância que o Projeto de Lei confere à proteção contra a discriminação e os vieses discriminatórios, consoante exposto na justificção constante do PL:

O projeto também reforça a proteção contra a discriminação, por meio de diversos instrumentos, como o direito à informação e compreensão, o direito à contestação, e em um direito específico de correção de vieses discriminatórios diretos, indiretos, ilegais ou abusivos, além das medidas de governança preventivas. Além de adotar definições sobre discriminação direta e indireta – incorporando, assim, definições da Convenção Interamericana contra o Racismo, promulgada em 2022 –, o texto tem como ponto de atenção grupos (hiper)vulneráveis tanto para a qualificação do que venha ser um sistema de alto risco como para o reforço de determinados direitos⁶⁵.

Conforme se observa, caso aprovado, o Projeto de Lei regulamentará todo o processo de desenvolvimento, implementação e uso de sistemas de inteligência artificial, os quais deverão respeitar diversos princípios e fundamentos, sendo conferida aos usuários uma grande variedade de direitos, visando especialmente à igualdade, à privacidade, à informação e à transparência de informações. Assim como no modelo europeu, os sistemas de inteligência artificial serão classificados conforme o risco, sendo vedados sistemas de risco excessivo. Ainda, estão dispostas previsões quanto à responsabilização e à aplicação de sanções por eventuais danos e atos ilícitos cometidos. Atualmente, o Projeto está sob análise das comissões permanentes do Senado, sem data para ser votado⁶⁶.

2.5 O ChatGPT

O ChatGPT, lançado em 30 de novembro de 2022, consiste em um modelo de

⁶⁵ SENADO FEDERAL. **Projeto de Lei nº 2338, de 2023**. Disponível em: <https://www25.senado.leg.br/web/atividade/materias/-/materia/157233>. Acesso em: 04 jul. 2023.

⁶⁶ SENADO FEDERAL. **Uso de inteligência artificial pelo poder público será sujeito a regulamentação**. 2023. Disponível em: <https://www12.senado.leg.br/noticias/materias/2023/05/12/uso-de-inteligencia-artificial-pelo-poder-publico-sera-sujeito-a-regulamentacao>. Acesso em: 04 jul. 2023. p. 31.

inteligência artificial generativa (*generative AI model*) desenvolvido pela empresa OpenAI, baseado em técnicas de processamento de linguagem natural, as quais envolvem o desenvolvimento de algoritmos e modelos que são capazes de entender e gerar linguagem humana⁶⁷. Segundo a desenvolvedora da aplicação⁶⁸, o ChatGPT é capaz de interagir de forma conversacional, de modo a permitir respostas a perguntas complementares, admitir erros, contestar premissas incorretas e rejeitar solicitações inadequadas.

A ferramenta possui diversas aplicabilidades como resumo de textos, elaboração de artigos, traduções, recomendações, cálculos matemáticos, sugestões de correção de programações, interpretação de imagens⁶⁹, geração de conteúdos como artigos jornalísticos e descrição de produtos⁷⁰. Ainda, pode ser aplicada em diversas áreas, como saúde, educação, serviços de atenção ao consumidor, criação de conteúdo, tradução de idiomas, entretenimento, finanças, ciência climática, *chatbots*, programação⁷¹.

Apesar das diversas possibilidades de utilização e do inegável avanço que representa no cenário tecnológico, ultimamente o ChatGPT vem despertando preocupação em especialistas, tendo em vista suas limitações e riscos. Dentre estes está o fato de que as respostas são construídas por meio da forma de interação com a ferramenta e da base de aplicação da plataforma, cuja última atualização ocorreu

⁶⁷ HARIRI, Walid. **Unlocking the Potential of ChatGPT: a comprehensive exploration of its applications, advantages, limitations, and future directions in natural language processing**. Annaba, Argélia, 2023. Disponível em: <https://doi.org/10.48550/arXiv.2304.02017>. Acesso em: 06 jul. 2023. p. 1.

⁶⁸ “We’ve trained a model called ChatGPT which interacts in a conversational way. The dialogue format makes it possible for ChatGPT to answer followup questions, admit its mistakes, challenge incorrect premises, and reject inappropriate requests.” (Treinamos um modelo chamado ChatGPT que interage de forma conversacional. O formato de diálogo permite que o ChatGPT responda a perguntas complementares, admita seus erros, conteste premissas incorretas e rejeite solicitações inadequadas.). OPENAI. **Introducing ChatGPT**. 2022. Disponível em: <https://openai.com/blog/chatgpt>. Acesso em: 05 jul. 2023 (tradução nossa).

⁶⁹ DUARTE, Filipe Ribeiro. **ChatGPT 4: os pontos de atenção e os riscos legais trazidos pelo uso da ferramenta**. 2023. Disponível em: <https://lawinnovation.com.br/chatgpt-4-os-pontos-de-atencao-e-os-riscos-legais-trazidos-pelo-uso-da-ferramenta>. Acesso em: 05 jul. 2023.

⁷⁰ HARIRI, Walid. **Unlocking the Potential of ChatGPT: a comprehensive exploration of its applications, advantages, limitations, and future directions in natural language processing**. Annaba, Argélia, 2023. Disponível em: <https://doi.org/10.48550/arXiv.2304.02017>. Acesso em: 06 jul. 2023. p. 1.

⁷¹ *Ibidem*, p. 7-12.

em setembro de 2021⁷². Walid Hariri⁷³ identifica riscos como: preconceito e padrão de linguagem ofensiva, habilidade limitada de compreensão de contexto, necessidade de grande quantidade de dados de treinamento, riscos de cibersegurança e falta de qualidade nas respostas.

Com efeito, o ChatGPT é capaz de ser fonte de desinformação e de preconceito, conforme reconhece a própria aplicação ao interagir com ela. Ao ser questionado se “O ChatGPT é capaz de fornecer informações falsas?”, responde o ChatGPT⁷⁴:

Sim, o ChatGPT pode fornecer informações falsas. Embora o modelo tenha sido treinado em uma ampla gama de dados e faça o melhor para fornecer respostas precisas e úteis, ele não tem a capacidade de verificar a veracidade de todas as informações disponíveis na internet. Portanto, em certas circunstâncias, o modelo pode gerar respostas que são imprecisas, desatualizadas ou mesmo falsas.

É importante ter cuidado ao confiar totalmente nas informações fornecidas pelo ChatGPT ou por qualquer outra fonte de inteligência artificial. Sempre verifique as informações por meio de fontes confiáveis e, se possível, consulte várias fontes para obter uma visão mais abrangente. A aplicação do pensamento crítico e o uso de discernimento são essenciais ao avaliar as informações provenientes de qualquer fonte, incluindo sistemas de IA como o ChatGPT.

Por sua vez, ao ser interrogado a respeito se “O ChatGPT pode proferir discurso de ódio e falas preconceituosas?”, refere:

O objetivo do ChatGPT é fornecer informações úteis e responder a perguntas de forma imparcial e neutra. O modelo foi treinado em uma ampla variedade de dados disponíveis publicamente, o que inclui a linguagem usada na internet. No entanto, medidas foram tomadas para mitigar a geração de discurso de ódio e falas preconceituosas durante o treinamento do modelo.

Apesar disso, é importante observar que o ChatGPT pode refletir os vieses e preconceitos presentes nos dados com os quais foi treinado. Isso significa que pode haver ocasiões em que o modelo produza respostas que contêm preconceitos ou expressões inadequadas. A OpenAI, a organização por trás do ChatGPT, está constantemente trabalhando para melhorar seus modelos e reduzir esses problemas. [...]

O referido risco de disseminação de discurso de preconceito e ofensivo está associado ao fato de que o ChatGPT é treinado com banco de dados incluindo fóruns

⁷² DUARTE, Filipe Ribeiro. **ChatGPT 4: os pontos de atenção e os riscos legais trazidos pelo uso da ferramenta**. 2023. Disponível em: <https://lawinnovation.com.br/chatgpt-4-os-pontos-de-atencao-e-os-riscos-legais-trazidos-pelo-uso-da-ferramenta>. Acesso em: 05 jul. 2023.

⁷³ HARIRI, Walid. **Unlocking the Potential of ChatGPT: a comprehensive exploration of its applications, advantages, limitations, and future directions in natural language processing**. Annaba, Argélia, 2023. Disponível em: <https://doi.org/10.48550/arXiv.2304.02017>. Acesso em: 06 jul. 2023. p. 13-15.

⁷⁴ Interação com a plataforma ChatGPT, disponível em: <https://chat.openai.com>. Acesso em: 06 jul. 2023.

on-line, posts de redes sociais e artigos jornalísticos, que podem conter linguagem preconceituosa e estereótipos ofensivos. Assim, a plataforma tem a capacidade de reforçar estereótipos ofensivos e perpetuar discriminação⁷⁵. Sustenta-se, inclusive, a dificuldade de afastamento de reprodução de informações nesse sentido, ainda que existentes ferramentas de revisão e de moderação das respostas geradas⁷⁶.

A propósito, em recente estudo⁷⁷ que visava à identificação de eventuais vieses do ChatGPT, constatou-se a existência de tendência político-ideológica à esquerda, ainda que a plataforma sustente neutralidade político-ideológica⁷⁸. A existência de vieses pode ser perigosa, na medida em que, sendo a plataforma utilizada por um grande número de pessoas, pode ser aproveitada para controle social, disseminação de desinformação, assim como manipulação de instituições e processos democráticos, representando obstáculo para a busca da verdade⁷⁹.

Sugere-se, pois, o desenvolvimento de pesquisas e métodos para detecção e mitigação de tendências. Ademais, David Rozado sustenta que é de grande relevância a apresentação aos usuários de argumentos equilibrados de todos os pontos de vistas legítimos sobre um determinado assunto⁸⁰.

3 A DIVULGAÇÃO DE *FAKE NEWS* E A PRÁTICA DE CRIMES DE ÓDIO: UMA ANÁLISE À LUZ DO DIREITO PENAL

3.1 A divulgação de *fake news* e o tratamento pelo direito penal brasileiro

A digitalização da produção e da divulgação de notícias vem desafiando a então

⁷⁵ HARIRI, Walid. **Unlocking the Potential of ChatGPT: a comprehensive exploration of its applications, advantages, limitations, and future directions in natural language processing**. Annaba, Argélia, 2023. Disponível em: <https://doi.org/10.48550/arXiv.2304.02017>. Acesso em: 06 jul. 2023. p. 13.

⁷⁶ DUARTE, Filipe Ribeiro. **ChatGPT 4: os pontos de atenção e os riscos legais trazidos pelo uso da ferramenta**. 2023. Disponível em: <https://lawinnovation.com.br/chatgpt-4-os-pontos-de-atencao-e-os-riscos-legais-trazidos-pelo-uso-da-ferramenta>. Acesso em: 05 jul. 2023.

⁷⁷ ROSADO, David. **The Political Biases of ChatGPT**. Hamilton, Nova Zelândia, 2023. Disponível em: <https://doi.org/10.3390/socsci12030148>. Acesso em: 06 jul. 2023. p. 1-8.

⁷⁸ Os resultados dos 15 testes de orientação política para ChatGPT foram consistentes entre os testes; 14 dos testes diagnosticaram as respostas do ChatGPT às suas perguntas como manifestando pontos de vista políticos de esquerda; veja a Figura 1. O teste restante (Teste de Nolan) diagnosticou as respostas do ChatGPT como politicamente centristas. (“The results of administering the 15 political orientation tests to ChatGPT were mostly consistent across tests; 14 of the tests diagnosed ChatGPT’s answers to their questions as manifesting left-leaning political viewpoints; see Figure 1. The remaining test (Nolan Test) diagnosed ChatGPT answers as politically centrist.”). ROSADO, David. **The Political Biases of ChatGPT**. Hamilton, Nova Zelândia, 2023. Disponível em: <https://doi.org/10.3390/socsci12030148>. Acesso em: 06 jul. 2023. p. 3.

⁷⁹ *Ibidem*, p. 5.

⁸⁰ *Ibidem*.

existente forma de produção e a conceituação do que efetivamente corresponde ao conteúdo da “notícia”, que pode assumir a dimensão de: (I) um relato de um evento recente, interessante e significativo; (II) um relato de eventos que afetam significativamente as pessoas; (III) um relato dramático de algo novo ou desviante⁸¹. As notícias são, em verdade, socialmente construídas, de modo que podem sofrer interferências externas.

Com efeito, as plataformas on-line passaram a fornecer espaço para que não-jornalistas possam atingir um público de massa. Segundo Edson C. Tandoc Jr., Zheng Wei Lim e Richard Ling⁸²:

A ascensão do “jornalismo cidadão” desafiou o vínculo entre notícias e jornalistas, pois não-jornalistas começaram a se envolver em atividades jornalísticas para produzir resultados jornalísticos, incluindo notícias [...]. Os jornalistas cidadãos foram inicialmente confinados aos blogs. Eventualmente, as redes sociais ofereceram uma plataforma mais ampla para não-jornalistas se envolverem no jornalismo [...]. Por meio de suas contas de redes sociais, os usuários podem postar informações, fotos, vídeos e narrativas sobre eventos interessantes que testemunham em primeira mão.

A partir desse contexto, informações passaram a ser divulgadas em rápida velocidade e em larga escala, alcançando grande número de pessoas em poucos minutos. Se, por um lado, a tecnologia vem contribuindo para a democratização do acesso informacional, por outro, vem trazendo sérios riscos quanto à qualidade, à confiabilidade e à veracidade do conteúdo divulgado.

Nessa esteira, surgem as chamadas *fake news* (notícias falsas), que, com as eleições presidenciais de 2016 nos Estados Unidos, marcadas por desinformação e inverdades, passou a ser um termo amplamente divulgado⁸³. O termo, além de ter sido escolhido como a palavra do ano em 2017 pelo *Collins English Dictionary*, com aumento de menções em 365% nesse ano, foi incluído no Dicionário, especialmente após a utilização por Donald Trump, que sistematicamente referia-se a notícias

⁸¹ TANDOC, Edson C.; LIM, Zheng Wei; LING, Richard. **Defining “Fake News”**. *Digital Journalism*, [S.L.], v. 6, n. 2, p. 137-153, 2017. Disponível em: <https://doi.org/10.1080/21670811.2017.1360143>. Acesso em: 22 jul. 2023. p. 140.

⁸² “The digitization of news has challenged traditional definitions of news. Online platforms provide space for non-journalists to reach a mass audience. The rise of citizen journalism challenged the link between news and journalists, as non-journalists began to engage in journalistic activities to produce journalistic outputs, including news [...]. Citizen journalists were initially confined to blogging. Eventually, social media offered a wider platform for non-journalists to engage in journalism [...] Through their social media accounts, users can post information, photos, videos, and narratives about newsworthy.”. *Ibidem*, p. 139 (tradução nossa).

⁸³ *Ibidem*, p. 147.

negativas sobre ele como *fake news*⁸⁴.

Segundo o *Collins English Dictionary*, o termo *fake news* corresponde a “informações falsas, muitas vezes sensacionalistas, divulgadas sob o disfarce de reportagens”⁸⁵. Por sua vez, o *Cambridge English Dictionary* também incluiu verbete, trazendo como definição “histórias falsas que aparentam ser notícias, divulgadas na internet ou em outros meios de comunicação, geralmente criadas para influenciar opiniões políticas ou como piada.”⁸⁶.

O termo *fake news*, compreendido como uma espécie do gênero “desinformação”⁸⁷, pode ser difícil de ser definido, assumindo a dimensão de um “termo guarda-chuva com múltiplas definições”⁸⁸. Uma revisão de estudos, a propósito, verificou a utilização do conceito em seis campos: (I) sátira de notícias, (II) paródia de notícias, (III) fabricação, (IV) manipulação, (V) publicidade e (VI) propaganda⁸⁹.

A título de exemplo, citam-se algumas definições encontradas:

I. Apresentação deliberada de alegações (normalmente) falsas ou enganosas como notícias, em que as alegações são enganosas de propósito⁹⁰.

II. Versão avançada e tecnológica da desinformação, com pedaços de

⁸⁴ BBC NEWS BRASIL. 'Fake News' é eleita palavra do ano e ganhará menção em dicionário britânico. 2017. Disponível em: <https://www.bbc.com/portuguese/internacional-41843695>. Acesso em: 23 jul. 2023.

⁸⁵ “Fake news: False, often sensational, information disseminated under the guise of news reporting.”. COLLINS ENGLISH DICTIONARY. **Definition of ‘fake news’**. Disponível em: <https://www.collinsdictionary.com/dictionary/english/fake-news>. Acesso em: 22 jul. 2023 (tradução nossa).

⁸⁶ “Fake news: False stories that appear to be news, spread on internet or using other media, usually created to influence political views or as a joke.”. CAMBRIDGE ENGLISH DICTIONARY. **Fake news**. Disponível em: <https://dictionary.cambridge.org/us/dictionary/english/fake-news>. Acesso em: 22 jul. 2023 (tradução nossa).

⁸⁷ GELFERT, Axel. **Fake News: a definition**. *Informal Logic*, [S.L.], v. 38, n. 1, p. 84-117, 2018. Disponível em: <http://doi.org/10.22329/il.v38i1.5068>. Acesso em: 22 jul. 2023. p. 103.

⁸⁸ LILLEKER, Darren G.. **Evidence to the Culture, Media and Sport Committee. 'Fake news' inquiry presented by the Faculty for Media & Communication**, Bournemouth University, 2017. Disponível em:

<https://eprints.bournemouth.ac.uk/28610/3/Evidence%20Submission%20-%20Fake%20News%20FINAL.pdf>. Acesso em: 23 jul. 2023.

⁸⁹ “A review of previous studies that have used the term fake news reveals six types of definition: (1) news satire, (2) news parody, (3) fabrication, (4) manipulation, (5) advertising, and (6) propaganda.”. TANDOC, Edson C.; LIM, Zheng Wei; LING, Richard. **Defining ‘Fake News’**. *Digital Journalism*, [S.L.], v. 6, n. 2, p. 137-153, 2017. Disponível em: <https://doi.org/10.1080/21670811.2017.1360143>. Acesso em: 22 jul. 2023. p. 147 (tradução nossa).

⁹⁰ “Fake news is the deliberate presentation of (typically) false or misleading claims as news, where the claims are misleading by design.”. GELFERT, Axel. **Fake News: a definition**. *Informal Logic*, [S.L.], v. 38, n. 1, p. 84-117, 2018. Disponível em: <http://doi.org/10.22329/il.v38i1.5068>. Acesso em: 22 jul. 2023. p. 108 (tradução nossa).

informações intencionalmente manipuladas que aparecem na internet e principalmente nas mídias sociais⁹¹.

III. Artigos de notícias intencional e comprovadamente falsos que podem enganar leitores⁹².

IV. Postagens virais baseadas em relatos fictícios feitas para parecer reportagens⁹³.

V. [...] Um fato distorcido, um boato, uma mentira até uma falsidade ou fraude, a depender da finalidade da situação⁹⁴.

Com efeito, a criação e a divulgação de *fake news* não são exatamente um fenômeno novo: o que é novidade é a possibilidade do uso massivo de sistemas digitais de comunicações com possibilidade de grande alcance por meio das redes sociais, que podem, em um ritmo acelerado, espalhar notícias e tornar qualquer mensagem viral⁹⁵. A divulgação de notícias falsas existe há muito tempo e os avanços tecnológicos vêm desencadeando um número maior de novas possibilidades de engano e fabricação⁹⁶.

⁹¹ “Accordingly, fake news could be understood as an advanced and technological version of misinformation, since with this term we mean pieces of intentionally manipulated information which appear on the Internet and in social media especially.” GIUSTI, Serena; PIRAS, Elisa. **Democracy and Fake News: Information Manipulation and Post-Truth Politics**. Routledge, [S.L.], 2021. *E-book*. Disponível em: <https://search.ebscohost.com/login.aspx?direct=true&AuthType=shib&db=nlebk&AN=2681917&lang=pt-br&scope=site&authtype=guest,shib&custid=s5837110&groupid=main&profile=eds>. Acesso em: 22 jul. 2023. p. 3.

⁹² “We define ‘fake news’ to be news articles that are intentionally and verifiably false, and could mislead readers.” ALLCOTT, Hunt; GENTZKOW, Matthew. **Social Media and Fake News in the 2016 Election**. Journal Of Economic Perspectives, [S.L.], v. 31, n. 2, p. 211-236, 2017. American Economic Association. Disponível em: <http://doi.org/10.1257/jep.31.2.211>. Acesso em: 23 jul. 2023. p. 213 (tradução nossa).

⁹³ “Contemporary discourse, particularly media coverage, seems to define fake news as referring to viral posts based on fictitious accounts made to look like news reports.” TANDOC, Edson C.; LIM, Zheng Wei; LING, Richard. **Defining “Fake News”**. Digital Journalism, [S.L.], v. 6, n. 2, p. 137-153, 2017. Disponível em: <https://doi.org/10.1080/21670811.2017.1360143>. Acesso em: 22 jul. 2023. p. 138 (tradução nossa).

⁹⁴ SILVA, Marco Antônio Marques da; BRAGA, Mariana Stuart Nogueira. **Fake news: é necessário novo tipo penal?**. Revista Paradigma, [S.L.], v. 30, n. 2, p. 152–174, 2022. Disponível em: <https://revistas.unaerp.br/paradigma/article/view/2433>. Acesso em: 22 jul. 2023. p. 158.

⁹⁵ GIUSTI, Serena; PIRAS, Elisa. **Democracy and Fake News: Information Manipulation and Post-Truth Politics**. Routledge, [S.L.], 2021. *E-book*. Disponível em: <https://search.ebscohost.com/login.aspx?direct=true&AuthType=shib&db=nlebk&AN=2681917&lang=pt-br&scope=site&authtype=guest,shib&custid=s5837110&groupid=main&profile=eds>. Acesso em: 22 jul. 2023. p. 3.

⁹⁶ “In short, the faking of news stories has been around for a long time, and every iteration of technological advancement, from the telegraph in the 19th century to contemporary social media algorithms, has unleashed new possibilities of deception and fabrication.” GELFERT, Axel. **Fake News: a definition**. Informal Logic, [S.L.], v. 38, n. 1, p. 84-117, 2018. Disponível em:

Nesse cenário, os sistemas de inteligência artificial, como veremos adiante, correspondem a avanços tecnológicos igualmente capazes de contribuir para a criação e divulgação das *fake news* em massa. A respeito dos riscos da IA na geração de notícias falsas sustenta Tom Taulli⁹⁷:

No curto prazo, contudo, a IA provavelmente terá maior impacto na guerra de informações; o que ainda pode ser destrutivo. Tivemos um vislumbre disso quando o governo russo interferiu nas eleições presidenciais americanas de 2016. A abordagem era de baixa tecnologia, uma vez que usava fazendas de redes sociais para disseminar notícias falsas, mas as consequências foram significativas.

Entende-se que são duas as motivações principais que fundamentam a produção de *fake news*: financeira e ideológica⁹⁸. Com relação à primeira motivação, tem-se que histórias falsas são capazes de se tornarem virais e, conseqüentemente, serem convertidas em receita publicitária. Por sua vez, com relação à motivação ideológica, verifica-se que notícias falsas podem ser produzidas para a promoção de ideias ou pessoas específicas, muitas vezes desacreditando outras.

As *fake news*, inicialmente conectadas com o cenário político, hoje passaram a tratar de todos os aspectos da vida. Assim, ao se apropriarem da aparência de notícia reais:

Escondem-se sob um verniz de legitimidade, pois assumem alguma forma de credibilidade ao tentar parecer notícias reais. Além disso, indo além da simples aparência de uma notícia, por meio do uso de bots de notícias, as notícias falsas imitam a onipresença das notícias ao construir uma rede de sites falsos⁹⁹.

A divulgação e a disseminação de *fake news*, além de passarem a representar riscos reputacionais à honra de indivíduos, correspondem, de igual sorte, a um risco à democracia, na medida em que podem estar envolvidas em contextos de embate e disputas ideológicas, nos quais as pessoas não questionam se a notícia é verdadeira ou falsa e se os fatos são embasados ou confiáveis. As *fake news*, conforme

<http://doi.org/10.22329/il.v38i1.5068>. Acesso em: 22 jul. 2023. p. 90 (tradução nossa).

⁹⁷ TAULLI, Tom. **Introdução à inteligência artificial**. São Paulo: Novatec, 2020. Tradução de: Luciana do Amaral Teixeira. p. 214.

⁹⁸ TANDOC, Edson C.; LIM, Zheng Wei; LING, Richard. **Defining “Fake News”**. Digital Journalism, [S.L.], v. 6, n. 2, p. 137-153, 2017. Disponível em: <https://doi.org/10.1080/21670811.2017.1360143>. Acesso em: 22 jul. 2023. p. 138 (tradução nossa).

⁹⁹ “Fake news hides under a veneer of legitimacy as it takes on some form of credibility by trying to appear like real news. Furthermore, going beyond the simple appearance of a news item, through the use of news bots, fake news imitates news’ omnipresence by building a network of fake sites.”. *Ibidem*, p. 147.

asseveram Marco Antônio Souza Alves e Emanuella Ribeiro Halfeld Maciel¹⁰⁰, passam a ser utilizadas “como armas no meio de um campo de batalha”, em que “a única coisa que importa é se a notícia favorece [determinada] posição em um contexto polarizado”.

Patrícia Campos Mello¹⁰¹, na mesma linha, trata a respeito dos riscos que a divulgações de informações falsas representa à democracia:

Na versão moderna do autoritarismo em que governantes não rasgam a Constituição nem dão golpes de Estado clássicos, mas corroem as instituições por dentro -, não é necessário censurar a internet. Nas “democracias iliberais”, segundo o vernáculo do primeiro-ministro húngaro Viktor Orbán, basta inundar as redes sociais e os grupos de WhatsApp com a versão dos fatos que se quer emplacar, para que ela se torne verdade - e abafe as outras narrativas, inclusive e sobretudo as reais.

Por essa razão, os notórios riscos e prejuízos das *fake news* – especialmente evidenciados e trazidos à tona nos recentes períodos eleitorais no Brasil – levaram à discussão acerca de como responsabilizar penalmente divulgadores e disseminadores de *fake news*. Surgiu, então, o questionamento acerca da possibilidade de subsunção da conduta aos tipos penais já existentes, ou se seria necessária a criação de novas leis e tipos penais.

Nesse sentido, passaram a tramitar no Congresso Nacional diversas iniciativas para criminalizar a divulgação de *fake news*, tendo em mente eventuais prejuízos econômicos, políticos e reputacionais que a propagação de informações falsas pode promover¹⁰². Com efeito, conforme sustentam Thamara Medeiros e Juliana Abrusio¹⁰³:

O debate sobre a criminalização das fake news não se limita à disseminação de notícias falsas, implica, em especial, nas possíveis consequências que a desinformação pode promover nas sociedades democráticas, ou seja, os prejuízos decorrentes das mentiras e dos boatos propagados por meio do ambiente virtual são inestimáveis e comumente irreversíveis.

A primeira iniciativa corresponde ao Projeto de Lei n.º 6.812/2017¹⁰⁴, de autoria

¹⁰⁰ ALVES, Marco Antônio Souza; MACIEL, Emanuella Ribeiro Halfeld. **O fenômeno das fake news: definição, combate e contexto**. Revista Internet & Sociedade, [S.L.], v. 1, n. 1, p. 144–171, 2020. Disponível em: <https://revista.internetlab.org.br/o-phenomeno-das-fake-news-definicao-combate-e-contexto>. Acesso em: 22 jul. 2023. p. 153.

¹⁰¹ MELLO, Patrícia Campos. **A máquina do ódio: notas de uma repórter sobre fake news e violência digital**. Companhia das Letras, São Paulo, 2020.

¹⁰² ABRUSIO, Juliana; Medeiros, Thamara. Os limites da criminalização da desinformação. In: (Coord.) RAIS, Diogo. **Fake News: a conexão entre a desinformação e o direito**. 3. ed. São Paulo: Thomson Reuters Brasil, 2022. *E-book*. Disponível em: <https://proview.thomsonreuters.com/launchapp/title/rt/monografias/156970600>. Acesso em: 23 jul. 2023. p. RB-15.3.

¹⁰³ *Ibidem*.

¹⁰⁴ CÂMARA DOS DEPUTADOS. **PL 6812/2017**. Disponível em: <https://www.camara.leg.br/proposicoesWeb/fichadetramitacao?idProposicao=2122678>. Acesso em: 23

do Deputado Federal Luiz Carlos Cauly (PSDB/PR), o qual propôs a tipificação da conduta de “divulgar ou compartilhar, por qualquer meio, na rede mundial de computadores, informação falsa ou prejudicialmente incompleta em detrimento de pessoa física ou jurídica”, prevendo pena de 2 a 8 meses de detenção e pagamento de 1.500 (mil e quinhentos) a 4.000 (quatro mil) dias-multa. Referido PL, atualmente, encontra-se apensado ao PL n.º 2630/2020, que institui a Lei Brasileira de Liberdade, Responsabilidade e Transparência na Internet.

Em segundo lugar, destaca-se o Projeto de Lei do Senado n.º 473/2017¹⁰⁵, em trâmite no Senado Federal, de autoria do Senador Ciro Nogueira (PP/PI), que dispôs acerca da inserção ao Código Penal do artigo 287-A, criando a figura típica da “divulgação de notícia falsa”. O PLS previa a seguinte disposição:

Divulgação de notícia falsa

Art. 287-A - Divulgar notícia que sabe ser falsa e que possa distorcer, alterar ou corromper a verdade sobre informações relacionadas à saúde, à segurança pública, à economia nacional, ao processo eleitoral ou que afetem interesse público relevante.

Pena – detenção, de seis meses a dois anos, e multa, se o fato não constitui crime mais grave.

§ 1º Se o agente pratica a conduta prevista no caput valendo-se da internet ou de outro meio que facilite a divulgação da notícia falsa:

Pena – reclusão, de um a três anos, e multa, se o fato não constitui crime mais grave.

§ 2º A pena aumenta-se de um a dois terços, se o agente divulga a notícia falsa visando a obtenção de vantagem para si ou para outrem.

Em terceiro lugar, menciona-se o PLS n.º 471/2018¹⁰⁶, de autoria do Senador Humberto Costa (PT/PE), o qual acresce ao Código Penal o artigo 288-B, correspondente ao crime de “criação ou divulgação ou divulgação de notícia falsa”, com redação de grande semelhança ao tipo penal do PLS n.º 473/2017, senão vejamos:

Criação ou divulgação de notícia falsa

Art. 288-B. Criar ou divulgar notícia que sabe ser falsa e para distorcer, alterar ou corromper gravemente a verdade sobre tema relacionado à saúde, à segurança pública, à economia nacional ou a outro interesse público relevante.

jul. 2023.

¹⁰⁵ SENADO FEDERAL. **Projeto de Lei do Senado nº 473, de 2017**. Disponível em: <https://www25.senado.leg.br/web/atividade/materias/-/materia/131758>. Acesso em: 23 jul. 2023.

¹⁰⁶ SENADO FEDERAL. **Projeto de Lei do Senado nº 471, de 2018**. Disponível em: <https://www25.senado.leg.br/web/atividade/materias/-/materia/134781>. Acesso em: 23 jul. 2023.

Pena – detenção, de seis meses a dois anos, e multa, se o fato não constitui crime mais grave.

§1º A pena é aumentada de um a dois terços, se o agente cria ou divulga a notícia falsa visando a obtenção de vantagem para si ou para outrem.

§ 2º É considerada notícia falsa, para os efeitos desta Lei, o texto não ficcional que, de forma intencional e deliberada, considerada a forma e as características da sua veiculação, tenha o potencial de ludibriar o receptor quanto à veracidade do fato.

§ 3º Não é considerada notícia falsa a manifestação de opinião, de expressão artística ou literária, ou o texto de conteúdo humorístico.

O PLS n.º 473/2017, além de prever novos artigos na Lei n.º 12.965/2014 (Marco Civil da Internet), cria o crime do artigo 354-B do Código Eleitoral, consistente na “criação ou divulgação de notícia falsa para afetar indevidamente o processo eleitoral”, assim redigido:

Criação ou divulgação de notícia falsa para afetar indevidamente o processo eleitoral

Art. 354-B. Criar ou divulgar notícia que sabe ser falsa para distorcer, alterar ou corromper gravemente a verdade relacionada ao processo eleitoral.

Pena – detenção, de seis meses a três anos, e multa, se o fato não constitui crime mais grave.

§1º A pena é aumentada de um a dois terços se o agente cria ou divulga a notícia falsa visando à obtenção de vantagem para si ou para outrem.

§ 2º É considerada notícia falsa, para os efeitos desta Lei, o texto não ficcional que, considerada a forma e características de sua veiculação, tenha o potencial de ludibriar o receptor quanto à veracidade do fato.

§ 3º Não é considerada notícia falsa a manifestação de opinião, de expressão artística ou literária, ou o texto de conteúdo humorístico.”

Em quarto lugar, destaca-se o PLS n.º 533/2018¹⁰⁷, de autoria do Senador Ataídes Oliveira (PSDB/TO), com redação exatamente igual à do PLS n.º 471/2018, o qual visa a alterar o Código Penal, o Código Eleitoral e o Marco Civil da Internet “para dispor sobre a definição das infrações penal, eleitoral e civil de criar ou divulgar notícia falsa, e cominar as respectivas penas”. Os Projetos de Lei do Senado n.º 473/2017, n.º 471/2018 e n.º 533/2018 foram arquivados, tendo em vista o final da legislatura, nos termos do § 1º do artigo 332 do Regimento Interno do Senado Federal¹⁰⁸.

¹⁰⁷ SENADO FEDERAL. **Projeto de Lei do Senado nº 533, de 2018**. Disponível em: <https://www25.senado.leg.br/web/atividade/materias/-/materia/134952>. Acesso em: 23 jul. 2023.

¹⁰⁸ “Art. 332. Ao final da legislatura serão arquivadas todas as proposições em tramitação no Senado, exceto: [...] § 1º Em qualquer das hipóteses dos incisos do caput, será automaticamente arquivada a proposição que se encontre em tramitação há duas legislaturas, salvo se requerida a continuidade de sua tramitação por 1/3 (um terço) dos Senadores, até 60 (sessenta) dias após o início da primeira sessão legislativa da legislatura seguinte ao arquivamento, e aprovado o seu desarquivamento pelo Plenário do Senado.” SENADO FEDERAL. **Regimento Interno do Senado Federal**. Disponível em:

Em quinto lugar, destaca-se o Projeto de Lei n.º 3.813/2021¹⁰⁹, apresentado pela Comissão Parlamentar de Inquérito da Pandemia, no qual, atualmente, aguarda-se a realização de audiência pública com o objetivo de instruir o Projeto de Lei. O PL cria o artigo 288-B do Código Penal, consistente no delito de “criação ou divulgação de notícia falsa”, com redação quase que semelhante à dos PLS n.º 471/2018 e PLS n.º 533/2018, à exceção de sutil alteração na redação do § 2º, que passa a considerar como notícia falsa “o texto, áudio, vídeo ou imagem não ficcional que, de modo intencional e deliberado, consideradas a forma e as características da sua veiculação, tenha o potencial de ludibriar o receptor quanto à veracidade do fato.”.

Da mesma forma, insere os parágrafos 4º e 5º à previsão do artigo 288-B do CP, definindo o delito de “notícia falsa sobre saúde pública”, com pena de 2 (dois) a 4 (quatro) anos de reclusão, e multa, se o fato não constitui crime mais grave. Além disso, insere no rol das medidas cautelares diversas da prisão do artigo 319 do Código de Processo Penal o inciso X, bem como o § 5º:

Art. 319. [...]

X - obrigação de promover a retirada de publicação em meios de comunicação, inclusive da rede mundial de computadores – internet e redes sociais, de notícia falsa que atente contra a saúde, a segurança, a economia ou outro interesse público relevante. [...]

§ 5º Aplicada a medida prevista no inciso X, o juiz mandará notificar os respectivos meios de comunicação ou provedores de aplicações de internet, dando-lhes ordem judicial contendo elementos que permitam a identificação específica do material a ser imediatamente retirado de publicação.

Os Projetos de Lei, contudo, foram objetos de crítica por setores da doutrina, que sustentam alguns problemas enfrentados, a saber:

I. Segundo Marco Antonio Marques da Silva e Marina Stuart Nogueira Braga¹¹⁰, não haveria bem jurídico constitucionalmente tutelável para inserir a modalidade das *fake news* como norma penal, na medida em que a mentira, apesar de imoral, não seria um bem jurídico constitucional apto à tutela penal. Na mesma linha, Thamara Medeiros e Juliana Abrusio¹¹¹ destacam a necessidade de se ter em mente o princípio

<https://www25.senado.leg.br/web/atividade/legislacao/regimento-interno>. Acesso em: 23 jul. 2023.

¹⁰⁹ SENADO FEDERAL. **Projeto de Lei nº 3813, de 2021**. Disponível em: <https://www25.senado.leg.br/web/atividade/materias/-/materia/150544>. Acesso em: 23 jul. 2023.

¹¹⁰ SILVA, Marco Antônio Marques da; BRAGA, Mariana Stuart Nogueira. **Fake news: é necessário novo tipo penal?**. Revista Paradigma, [S.L.], v. 30, n. 2, p. 152–174, 2022. Disponível em: <https://revistas.unaerp.br/paradigma/article/view/2433>. Acesso em: 22 jul. 2023. p. 163.

¹¹¹ ABRUSIO, Juliana; MEDEIROS, Thamara. Os limites da criminalização da desinformação. In: (Coord.) RAIS, Diogo. **Fake News: a conexão entre a desinformação e o direito**. 3. ed. São Paulo: Thomson Reuters Brasil, 2022. E-book. Disponível em:

da *ultima ratio* do direito penal, refletindo acerca da possibilidade de que a intervenção criminal garanta segurança sem comprometer os princípios fundamentais do direito penal de um Estado democrático¹¹².

II. Os verbos dos tipos penais – distorcer, alterar ou corromper a verdade – seriam amplos e vagos, de modo que comprometeriam a tipicidade, sendo de difícil verificação e mensuração as ações de divulgação e compartilhamento¹¹³.

III. A identificação do autor da disseminação de *fake news* seria algo difícil, haja vista que “a cadeia de compartilhamento é complexa, o que comprometeria a individualização do sujeito ativo”¹¹⁴.

IV. A valoração do dolo da conduta e dos elementos normativos do tipo relacionados ao conhecimento da inverdade/falsidade da notícia transitaria na órbita da imprecisão¹¹⁵. A propósito, Marco Antônio Sousa Alves e Emanuella Ribeiro Halfed Maciel¹¹⁶ sustentam que as *fake news* não envolvem necessariamente uma ação dolosa, um ato consciente de deturpação da realidade e de enganação, tratando-se de uma ação para corroborar ou enfraquecer determinada narrativa.

Nesse sentido, alguns doutrinadores entendem que outros ramos do direito – como o direito civil ou o direito administrativo – seriam igualmente aptos a dar uma sanção adequada à divulgação de *fake news*, sendo desnecessária a criação de novas figuras típicas¹¹⁷. Marco Antonio Marques da Silva e Marina Stuart Nogueira¹¹⁸, nessa linha, sustentam:

A saber, em uma situação em que notícia falsa é publicada no âmbito eleitoral, poderia o direito eleitoral prever sanções das mais variadas, chegando a eventual inelegibilidade se proporcional for ao ato vislumbrado. Outras situações que podem afetar direitos da personalidade podem ser passíveis de avaliação sob a ótica do direito civil, com reparação de danos, obrigações de fazer, deixar de fazer, e especificando alguma retratação se necessário. Evidentemente, não é só o direito penal essencial para salvaguardar fatos

<https://proview.thomsonreuters.com/launchapp/title/rt/monografias/156970600>. Acesso em: 23 jul. 2023. p. RB-15.3.

¹¹² *Ibidem*.

¹¹³ *Ibidem*.

¹¹⁴ *Ibidem*.

¹¹⁵ *Ibidem*.

¹¹⁶ ALVES, Marco Antônio Souza; MACIEL, Emanuella Ribeiro Halfed. **O fenômeno das fake news: definição, combate e contexto**. Revista Internet & Sociedade, [S.L.], v. 1, n. 1, p. 144–171, 2020. Disponível em: <https://revista.internetlab.org.br/o-fenomeno-das-fake-news-definicao-combate-e-contexto>. Acesso em: 22 jul. 2023. p. 153.

¹¹⁷ SILVA, Marco Antônio Marques; BRAGA, Mariana Stuart Nogueira. **Fake news: é necessário novo tipo penal?**. Revista Paradigma, [S.L.], v. 30, n. 2, p. 152–174, 2022. Disponível em: <https://revistas.unaerp.br/paradigma/article/view/2433>. Acesso em: 22 jul. 2023. p. 171.

¹¹⁸ *Ibidem*, p. 167.

socialmente relevantes. Outros ramos do direito, como o direito civil ou o administrativo são, por vezes, mais eficazes e proporcionais às condutas que se visa tutelar.

Além disso, entende-se que, considerando que as *fake news* podem atingir bens jurídicos existentes, seria possível a utilização de um tipo penal já previsto no Código Penal ou no Código Eleitoral. Desta forma:

Caso uma notícia falsa traga afirmações errôneas que lesionem a honra de determinada pessoa ou grupo de pessoas, para isso existem os crimes contra a honra (artigo 138 e seguintes do Código Penal) e até mesmo lei penal que visa punir crimes de preconceito (Lei n. 7.716/1989). Ademais, para fatos que gerem prejuízo econômico envolvendo informação falsa como meio de fraude, existe o crime de estelionato (artigo 171 do Código Penal). Quando há finalidade eleitoral, não se pode deixar de considerar o artigo 323, do Código Eleitoral que visa criminalizar a divulgação de fatos inverídicos em ambiente eleitoral.

Com efeito, concordamos com a dificuldade na definição da autoria e da aferição da tipicidade objetiva e subjetiva da divulgação, especialmente considerando a velocidade e a quantidade de informações e notícias compartilhadas diariamente. Da mesma forma, salutar a observância aos princípios limitadores do direito penal, como o princípio da *ultima ratio* do direito penal.

Entendemos que, de fato, a conduta da divulgação de notícias/informações falsas já estaria abarcada por alguns tipos penais, especialmente se atingirem a honra de um indivíduo, a exemplo dos crimes de calúnia (artigo 138 do CP), difamação (artigo 139 do CP) e injúria (artigo 140 do CP). Por essa razão, ainda que, até o momento, não exista tipo penal específico concernente à divulgação de *fake news*, seria possível se falar em responsabilização penal ao se enquadrar a conduta em tipos penais já existentes.

Em que pese a compreensão de eventuais problemas, entendemos necessária a existência de um tipo penal específico que trate da conduta de divulgação de notícia falsa. Isso porque a divulgação de *fake news* pode alcançar uma grande dimensão, produzindo resultados altamente danosos, de modo que se estaria tutelando, em verdade, o próprio Estado Democrático de Direito. Com efeito, a pandemia de Covid-19 trouxe à tona a capacidade de o ser humano criar narrativas e inverdades para instrumentalizar, de encontro ao conhecimento científico, visões político-ideológicas, prejudicando, inclusive, a vida e a saúde coletivas de forma intencional.

Do mesmo modo, os recentes processos eleitorais de 2018, 2020 e 2022

evidenciaram a instrumentalização de *fake news* como “arma”. Nesse cenário em que inverdades deixam de ser contestadas e passam a repercutir como verdades, cria-se um cenário cada vez mais polarizado, culminando com ações violentas e antidemocráticas, a exemplo dos recentes ocorridos no dia 08 de janeiro de 2023, no qual manifestantes apoiadores do ex-presidente Jair Bolsonaro invadiram e depredaram o Congresso Nacional, o Supremo Tribunal Federal (STF) e o Palácio do Planalto, em Brasília¹¹⁹.

3.2 Os crimes de ódio em tempos de pós-modernidade

Dentre todas as transformações que a internet e o universo digital propiciaram está o surgimento de uma nova criminalidade, com o aparecimento de novas figuras típicas, bem como de novos mecanismos e instrumentos de produção de crimes. Nesse sentido, conforme sustenta Marcelo Xavier de Freitas Crespo¹²⁰, o ciberespaço é “campo para o cometimento de delitos que já são tipificados em ordenamentos jurídicos, mas, também, é área onde condutas ainda não necessariamente incriminadas [...], mas altamente danosas, ocorrem.”. Surgem, então, os chamados “crimes digitais”:

Concretamente, os crimes digitais importam nas menções às condutas de acesso não autorizado a sistemas informáticos, ações destrutivas nesses sistemas, a interceptação de comunicações, modificações de dados, infrações a direitos de autor, incitação ao ódio e discriminação, escárnio religioso, difusão de pornografia infantil, terrorismo, entre outros¹²¹.

Além da possibilidade de prática de crimes contra os próprios sistemas informáticos, é possível a perpetração de condutas delituosas tradicionais por meio do universo digital. Isso porque o ciberespaço possui algumas características que podem contribuir para o desenvolvimento de espaço propício para criminalidade, a saber¹²²:

I. Capacidade de processar, armazenar e circular, de forma automatizada e em tempo real, grandes quantidades de informações nos mais variados formatos digitais.

II. Estrutura propícia para dar repercussão e amplo alcance a publicações e

¹¹⁹ G1 GLOBO. **Terrorismo em Brasília: o dia em que bolsonaristas criminosos depredaram Planalto, Congresso e STF**. 2023. Disponível em: <https://g1.globo.com/df/distrito-federal/noticia/2023/01/08/o-dia-em-que-bolsonaristas-invadiram-o-congresso-o-planalto-e-o-stf-como-isso-aconteceu-e-quais-as-consequencias.ghtml>. Acesso em: 24 jul. 2023.

¹²⁰ CRESPO, Marcos Xavier de Freitas. **Crimes digitais**. São Paulo: Saraiva, 2011. *E-book*. Disponível em: <https://app.minhabiblioteca.com.br/#/books/9788502136663>. Acesso em: 25 jul. 2023. p. 20.

¹²¹ *Ibidem*, p. 20.

¹²² *Ibidem*.

postagens, com a criação de fóruns, perfis e páginas em redes sociais.

III. Estrutura descentralizada e não hierarquizada da internet, inviabilizando a existência de órgãos de controle da informação circulante.

IV. Grande número de usuários, que frequente e livremente acessam e difundem informações, “de modo que os internautas passam a ser potenciais vítimas, mas também potenciais sujeitos ativos de delitos”.

V. Possibilidade de acesso de forma ilegítima às tecnologias, tendo em vista suas características físicas e técnicas.

A facilidade em se manifestar, de maneira aparentemente irrestrita e ilimitada, pode levar à conclusão de que a internet é “uma terra sem dono”, sendo a liberdade de manifestação do pensamento – prevista no artigo 5º, inciso IV, da Constituição Federal de 1988¹²³ – e a liberdade de expressão – prevista no artigo 5º, inciso IX, da CF/88¹²⁴ – vislumbradas como direitos amplos e escudos para manifestações ofensivas e para o discurso de ódio. Questiona-se, pois, se a liberdade de expressão, de comunicação e de pensamento são, de fato, absolutas, irrestritas e ilimitadas.

Conforme assevera Eduardo Carlos Bianca Bittar¹²⁵, o direito brasileiro não concebe a amplitude da liberdade de expressão do direito norte-americano, que “convive de forma liberal com atos e expressões de ódio (*hate speech*)”. Em verdade, o direito brasileiro segue a tendência do direito alemão, conferindo “prevalência à dignidade humana e à proteção da reputação e da honra”¹²⁶. Dessa maneira, tem-se que:

[...] O direito brasileiro não protege a vontade de oprimir de uns que poderiam submeter discursivamente (e agressivamente) outros. O abuso, a violência, o racismo e a discriminação não podem se travestir sob o *nomem iuris* da liberdade para permitir a circulação da agressão e da violência verbal, especialmente nos meios digitais, em função de sua propagação, de seu

¹²³ Art. 5º. [...] IV - é livre a manifestação do pensamento, sendo vedado o anonimato;

¹²⁴ Art. 5º. [...] IX - é livre a expressão da atividade intelectual, artística, científica e de comunicação, independentemente de censura ou licença;

¹²⁵ BITTAR, Eduardo Carlos Bianca. O Direito na Era Digital: responsabilidade civil e penal pelo uso indevido das redes sociais. In: SARLET, Ingo Wolfgang; SARLET, Gabrielle Bezerra Sales; BITTAR, Eduardo Carlos Bianca. **Inteligência artificial, proteção de dados pessoais e responsabilidade na era digital**. São Paulo: ExpressaJur, 2022. E-book. Disponível em: <https://app.minhabiblioteca.com.br/#/books/9786555599527>. Acesso em: 25 jul. 2023. p. 22.

¹²⁶ *Ibidem*.

alcance, de sua permanência e de sua capilaridade¹²⁷.

Com efeito, segundo a criminologista canadense Barbara Perry¹²⁸, “crimes de ódio” (“*hate crimes*”) são crimes que envolvem:

[...] atos de violência e intimidação e com frequência são direcionados a grupos que já são estigmatizados e marginalizados. Como tal, é um mecanismo de poder e de opressão, cuja intenção é reafirmar as hierarquias precárias que caracterizam uma ordem social dada. Ele tenta reproduzir, ao mesmo tempo, a hegemonia ameaçada (real ou imaginária) do grupo do autor e a identidade subordinada “adequada” do grupo da vítima. É uma forma de marcar tanto o Eu quanto o Outro de tal forma a restabelecer suas posições relativas “adequadas”, conforme são dadas e reproduzidas por ideologias e padrões mais abrangentes de desigualdade social e política.

James B. Jacobs e Kimberly Potter¹²⁹, por sua vez, sustentam que os crimes de ódio não se referem necessariamente ao “ódio”, mas, sim, à discriminação e ao preconceito, sendo um conceito socialmente construído, que pode abranger um grande número de ofensores e de condutas. A expressão “crime de ódio” foi cunhada no final dos anos 1980 para enfatizar condutas criminais motivadas pelo preconceito, que pode representar um conceito complicado, amplo e nebuloso, na medida em que o preconceito pode se manifestar de diferentes formas.

O preconceito, por sua vez, na linha do que refere Christiano Jorge Santos¹³⁰, consiste na formulação de ideias amparadas em concepções prévias que não foram objeto de devida reflexão ou que foram desenvolvidas por ideias deturpadas, sendo, portanto, um “pré-conceito”. Pode o preconceito se manifestar por meio de diferentes maneiras, seja pela fala, seja pelo gesto, seja por escrito.

Exterioriza-se preconceito, também, por atos de teor discriminatório. A discriminação, a seu turno, corresponde a uma:

Segregação (negativa) dolosa, comissiva ou omissiva, adotada contra alguém por pertencer, real ou supostamente, a uma raça, cor, etnia, religião ou por conta de sua procedência nacional e que visa a atrapalhar, limitar ou tolher o exercício regular do direito da pessoa discriminada, contrariando o princípio

¹²⁷ *Ibidem*.

¹²⁸ PERRY, Barbara. **In the Name of Hate: Understanding Hate Crimes**. Routledge: Nova York, 2001. *E-book*. Disponível em: <https://doi.org/10.4324/9780203905135>. Acesso em: 25 jul. 2023. p. 10.

¹²⁹ JACOBS, James B; POTTER, Kimberly. **Hate crimes: criminal law & identity politics**. Oxford University Press: Nova York, 2001. *E-book*. Disponível em: <https://search.ebscohost.com/login.aspx?direct=true&AuthType=shib&db=nlebk&AN=146872&lang=pt-br&scope=site&authType=guest,shib&custid=s5837110&groupid=main&profile=eds>. Acesso em: 25 jul. 2023. p. 11-28.

¹³⁰ SANTOS, Christiano Jorge. **Crimes de preconceito e de discriminação**. 2. ed. São Paulo: Saraiva, 2010. *E-book*. Disponível em: <https://app.minhabiblioteca.com.br/#/books/9788502113114>. Acesso em: 25 jul. 2023. p. 43.

constitucional da isonomia¹³¹.

Com o discurso de ódio, visa-se, por meio do uso indevido da fala, de gestos ou da escrita, a afetar a honra e a personalidade de outrem, causando danos e abalos e atingindo indivíduos ou grupo de pessoas, geralmente em razão da raça, gênero, etnia, religião, ideologia ou orientação sexual¹³². O discurso de ódio, nesse sentido, afeta tanto a sociedade, haja vista que cria um ambiente social de alta intolerância (um *hostile environment*), quanto o indivíduo enquanto pessoa, considerando que atinge a esfera de dignidade, sentindo-se o indivíduo humilhado e desvalorizado¹³³. Assim, o discurso do ódio, segundo Winfried Brugger¹³⁴, refere-se à utilização de palavras “que tendem a insultar, intimidar ou assediar pessoas em virtude de sua raça, cor, etnicidade, nacionalidade, sexo ou religião ou que têm a capacidade de instigar violência, ódio ou discriminação contra tais pessoas.”.

Irineu Barreto¹³⁵, a propósito, associa as *fake news* ao discurso de ódio, sustentado o “caráter indissociável” de sua relação. Isso porque as *fake news* são potencializadas pela sua associação com um contexto social mais amplo e no ambiente da internet, especialmente considerando o efeito dos algoritmos das redes sociais, plataformas de vídeos e buscadores. Na interação entre tecnologia e humanos, o discurso de ódio possui a capacidade de ganhar maior engajamento, exercendo grande poder de atração sobre os usuários da internet. Dessa forma:

A instigação do ódio é o componente mais importante para disseminação da Fake News. O discurso que mais adere e se propaga é aquele que elege inimigos – ideológicos, políticos, morais, identitários, de classe social – servindo simultaneamente como conteúdo e força de propulsão do Discurso de ódio¹³⁶.

Nesse cenário, o direito brasileiro prevê um conjunto de normativas e

¹³¹ *Ibidem*, p. 46.

¹³² BITTAR, Eduardo Carlos Bianca. O Direito na Era Digital: responsabilidade civil e penal pelo uso indevido das redes sociais. In: SARLET, Ingo Wolfgang; SARLET, Gabrielle Bezerra Sales; BITTAR, Eduardo Carlos Bianca. **Inteligência artificial, proteção de dados pessoais e responsabilidade na era digital**. São Paulo: ExpressaJur, 2022. *E-book*. Disponível em: <https://app.minhabiblioteca.com.br/#/books/9786555599527>. Acesso em: 25 jul. 2023. p. 22.

¹³³ *Ibidem*, p. 22.

¹³⁴ BRUGGER, Winfried. **Proibição ou proteção do discurso do ódio?: algumas observações sobre o direito alemão e o americano**. Direito Público, Porto Alegre, v. 4, n. 15, 2007. Disponível em: <https://www.portaldeperiodicos.idp.edu.br/direitopublico/article/view/1418>. Acesso em: jul. 2023. p. 118.

¹³⁵ BARRETO, Irineu. **Fake News: anatomia da desinformação, discurso de ódio e erosão da democracia**. São Paulo: ExpressaJur, 2022. *E-book*. Disponível em: <https://app.minhabiblioteca.com.br/reader/books/9786555598841>. Acesso em: 28 jul. 2023. p. 14.

¹³⁶ *Ibidem*, p. 15.

dispositivos que enfrentam o discurso e o crime de ódio, tanto na esfera constitucional, quanto no âmbito criminal. Primeiramente, cita-se os incisos XLI e XLII do artigo 5º da CF/88, os quais preveem que “a lei punirá qualquer discriminação atentatória dos direitos e liberdades fundamentais”, bem como que “a prática do racismo constitui crime inafiançável e imprescritível, sujeito à pena de reclusão, nos termos da lei”.

Em segundo lugar, há a Lei n.º 7.716/89, a qual prevê, em seu artigo 1º, que serão punidos os crimes resultantes de discriminação ou preconceito de raça, cor, etnia, religião ou procedência. O Supremo Tribunal Federal, na Ação Direta de Inconstitucionalidade por Omissão n.º 26, julgada em 13/06/2019¹³⁷, em razão da inércia e omissão legislativa, incluiu a homofobia e a transfobia entre as modalidades de condutas assinaladas pela Lei n.º 7.716/89, entendendo que se traduzem em expressões de racismo, senão vejamos:

Até que sobrevenha lei emanada do Congresso Nacional destinada a implementar os mandados de criminalização definidos nos incisos XLI e XLII do art. 5º da Constituição da República, as condutas homofóbicas e transfóbicas, reais ou supostas, que envolvem aversão odiosa à orientação sexual ou à identidade de gênero de alguém, por traduzirem expressões de racismo, compreendido este em sua dimensão social, ajustam-se, por identidade de razão e mediante adequação típica, aos preceitos primários de incriminação definidos na Lei nº 7.716, de 08/01/1989, constituindo, também, na hipótese de homicídio doloso, circunstância que o qualifica, por configurar motivo torpe (Código Penal, art. 121, § 2º, I, “in fine”)¹³⁸.

A prática de crimes de ódio, por sua vez, adquire distinto relevo no cenário digital, na medida em que “as plataformas digitais consentem alcance mundial, visibilidade e permanência a palavras, gestos, imagens, sons e vídeos, o que cria uma grave equação, quando se trata de conferir amplificação ao *hate speech*”¹³⁹. O ambiente virtual, como destacado por Magistrado da 22ª Vara Federal de Porto Alegre, RS, ao proferir sentença condenatória de sujeito que divulgou publicações de cunho nazista na rede social Twitter, apresenta-se “como um espaço atrativo e fomentador

¹³⁷ BRASIL. Supremo Tribunal Federal. **Ação Direta de Inconstitucionalidade por Omissão n.º 26**. Requerente: Partido Popular Socialista. Relator: Ministro Celso de Mello. Tribunal Pleno. Julgado em 13/06/2019. Disponível em: <https://redir.stf.jus.br/paginadorpub/paginador.jsp?docTP=TP&docID=754019240>. Acesso em: 25 jul. 2023.

¹³⁸ *Ibidem*, p. 5.

¹³⁹ BITTAR, Eduardo Carlos Bianca. O Direito na Era Digital: responsabilidade civil e penal pelo uso indevido das redes sociais. In: SARLET, Ingo Wolfgang; SARLET, Gabrielle Bezerra Sales; BITTAR, Eduardo Carlos Bianca. **Inteligência artificial, proteção de dados pessoais e responsabilidade na era digital**. São Paulo: ExpressaJur, 2022. E-book. Disponível em: <https://app.minhabiblioteca.com.br/#!/books/9786555599527>. Acesso em: 25 jul. 2023. p. 23.

do discurso de ódio, pois a sensação de anonimato e a distância entre os interlocutores se potencializam nas redes sociais”¹⁴⁰.

Por essa razão, tem-se verificado um movimento das plataformas de tecnologia ao pretenderem regular e definir balizas para o tipo de conteúdo que pode ser publicado, restringindo a possibilidade de divulgação de discurso de ódio, inclusive sujeita a penalidades. A título de exemplo, nas “Políticas do YouTube”, na sessão “Política de discurso de ódio”¹⁴¹, consta que:

O discurso de ódio não é permitido no YouTube. Removeremos todo conteúdo que promova a violência ou o ódio contra indivíduos ou grupos com base em qualquer uma das seguintes características: Idade; Classe social; Deficiência; Etnia; Identidade e expressão de gênero; Nacionalidade; Raça; Situação de imigração; Religião; Sexo/gênero; Orientação sexual; Vítimas de um evento violento em grande escala e os familiares dessas pessoas; Veteranos de guerra. [...]

Não bastasse, o direito brasileiro igualmente parece estar se atentando às particularidades e à gravidade do discurso de ódio produzido no cenário digital. Cita-se, a propósito, a recente nova redação do § 2º do artigo 20 da Lei n.º 7.716/89, introduzida com a Lei n.º 14.532/2023, que passou a prever pena de reclusão de dois a cinco anos e multa – superior à prevista no caput (reclusão de um a três anos e multa) – pela prática, induzimento ou incitação à discriminação ou ao preconceito de raça, cor, etnia, religião ou procedência nacional cometidos por intermédio dos meios de comunicação social, de publicação em redes sociais, da rede mundial de computadores ou de publicação de qualquer natureza.

Nessa linha, verifica-se que o crescente aumento na utilização de sistemas de inteligência artificial impõe novos desafios e questionamentos no que tange aos crimes de ódio praticados no meio digital. Até pouco tempo, considerava-se que o ambiente virtual era apenas instrumento e cenário para prática de crimes. Hoje, com as novas tecnologias desenvolvidas, passa-se, inclusive, a verificar casos em que os próprios sistemas de inteligência artificial, por sua atividade própria, são capazes de produzir discursos ofensivos, preconceituosos e discriminatórios, conforme veremos a seguir.

¹⁴⁰ Número do processo não divulgado. CONSULTOR JURÍDICO. **Publicações de cunho nazista no Twitter configuram crime de racismo, diz juiz**. 2023. Disponível em: <https://www.conjur.com.br/2023-jul-17/publicacoes-cunho-nazista-twitter-configuram-crime-racismo>. Acesso em: 25 jul. 2023.

¹⁴¹ GOOGLE SUPPORT. **Política de discurso de ódio**. Disponível em: <https://support.google.com/youtube/answer/2801939?hl=pt-BR>. 2023. Acesso em: 25 jul. 2023.

3.3 Os sistemas de inteligência artificial como divulgadores de *fake news* e praticantes de crimes de ódio

“Inteligência artificial” é, sem dúvida, um dos principais temas do momento. Ao passo que se tem destacado os diversos benefícios que podem representar nas relações humanas, acelerando e otimizando processos e tarefas, tem-se discutido os impactos negativos e malefícios que os sistemas de inteligência artificial podem trazer. Atualmente, são vários os episódios relatados de eventuais excessos e aplicações desvirtuadas envolvendo a tecnologia.

No que tange à divulgação de *fake news*, vem-se apontando que o ChatGPT será a mais poderosa ferramenta para espalhar desinformação que já existiu na internet, visto que uma narrativa falsa pode ser feita em maior escala e com maior frequência. Além disso, os sistemas de inteligência artificial têm tornado cada vez mais fácil e barato produzir desinformação, com possibilidade de divulgação de teorias da conspiração de maneiras confiáveis e persuasivas. Assim, é possível a redução de erros humanos, evidenciados em *fake news* manualmente criadas, como erros de sintaxe e de tradução¹⁴².

Um recente estudo, publicado em junho de 2023, conduzido por pesquisadores da Universidade de Zurique (Suíça), demonstrou que a maioria dos 697 participantes do experimento, que foram expostos a *tweets* noticiosos escritos por humanos e pelo ChatGPT, foram incapazes de distinguir se as informações eram escritas pelo sistema de IA ou por humano¹⁴³. Da mesma maneira, as *fake news* criadas pelo ChatGPT foram consideradas como verdadeiras em maior frequência. Segundo referem os pesquisadores Giovanni Spitale, Nikola Biller-Adorno e Federico Germani:

Os resultados do nosso estudo, incluindo 697 participantes, mostram que o GPT-3 é uma faca de dois gumes: em comparação com humanos, pode produzir informação precisa que é mais fácil de entender, mas, também, pode produzir desinformação mais convincente. Mostramos que os humanos não conseguem distinguir entre *tweets* gerados pelo GPT-3 e escritos por reais usuários do Twitter¹⁴⁴.

¹⁴² FOLHA DE SÃO PAULO. **ChatGPT será maior espalhador de desinformação que já existiu, diz pesquisador**. 2023. Disponível em: <https://www1.folha.uol.com.br/tec/2023/02/chatgpt-sera-maior-espalhador-de-desinformacao-que-ja-existiu-diz-pesquisador.shtml>. Acesso em: 26 jul. 2023.

¹⁴³ SPITALE, Giovanni; BILLER-ANDORNO, Nikola; GERMANI, Federico. **AI model GPT-3 (dis)inform us better than humans**. *Science Advances*: v. 9, n. 26. 2023. Disponível em: <https://www.science.org/doi/10.1126/sciadv.adh1850>. Acesso em: 04 jul. 2023. p. 1.

¹⁴⁴ “The results of our preregistered study, including 697 participants, show that GPT-3 is a doubleedge

No mesmo estudo, destacou-se que “informações produzidas pelo GPT-3 não só são mais efetivas em informar e desinformar humanos, como também o fazem de maneira mais eficiente em menos tempo.”¹⁴⁵. Concluiu-se que “*tweets* produzidos pelo GPT-3 podem simultaneamente informar e desinformar melhor do que *tweets* orgânicos”¹⁴⁶.

O fato de a divulgação de informação ser mais eficiente pelo sistema pode estar ligado à possibilidade de geração de texto mais fácil de ler e entender se comparado com textos escritos por humanos. Faz-se, pois, o alerta:

Embora o GPT-3 possa ser uma ótima ferramenta para traduções automáticas, classificação de texto, sistemas de diálogo/chatbot, resumo de conhecimento, resposta a perguntas, escrita criativa [...], detecção de discurso de ódio [...] e escrita automática de código [...], também pode ser usado para produzir desinformação, spam, phishing, abusar de processos legais e governamentais, redigir ensaios acadêmicos fraudulentos e pretexto de engenharia social¹⁴⁷.

Outrossim, ainda notícias falsas mal escritas podem ser extremamente influentes, sendo que, muitas vezes, verifica-se que quanto menos coerente e lógica for uma história falsa, maior a probabilidade de ela se tornar viral. Noah Gianciracusa¹⁴⁸ sustenta que a possibilidade de o ChatGPT propagar *fake news* no futuro dependerá de o quanto o ChatGPT fornece uma maneira mais barata e rápida de produzir *fake news* se comparado com seres humanos e o custo que a desenvolvedora cobrará dos usuários.

Nesse sentido, inclusive pesquisadores da desenvolvedora da plataforma ChatGPT, OpenAI, em artigo publicado em 2019, destacaram “preocupação de que seus recursos possam reduzir os custos de campanha de desinformação”¹⁴⁹ e ajudar

sword: In comparison with humans, it can produce accurate information that is easier to understand, but it can also produce more compelling disinformation. We also show that humans cannot distinguish between tweets generated by GPT-3 and written by real Twitter users.” *Ibidem*, p. 1 (tradução nossa).

¹⁴⁵ “[...] information produced by GPT-3 is not only more effective to inform and disinform humans but also does so more efficiently in less time.”. *Ibidem*, p. 3 (tradução nossa).

¹⁴⁶ “[...] tweets produced by GPT-3 can both inform and disinform better than organic tweets.”. *Ibidem*, p. 4 (tradução nossa).

¹⁴⁷ “While GPT-3 can be a great tool for machine translations, text classification, dialogue/chatbot systems, knowledge summarizing, question answering, creative writing (2, 5, 6), detecting hate speech (7), and automatic code writing (2, 8), it can also be used to produce “misinformation, spam, phishing, abuse of legal and governmental processes, fraudulent academic essay writing, and social engineering pretexting” (1, 9–11).” *Ibidem*, p 1.

¹⁴⁸ GIANCIRACUSA, Noah. **How Algorithms Create and Prevent Fake News**. Nova York: Springer Science Business Medida, 2021. *E-book*. Disponível em: <https://doi.org/10.1007/978-1-4842-7155-1> Acesso em: 25 jul. 2023. p. 29.

¹⁴⁹ “In our initial post on GPT-2, we noted our concern that its capabilities could lower costs of

na busca maliciosa “de ganho monetário, uma agenda política específica e/ou um desejo de criar caos ou confusão”¹⁵⁰. A propósito, vem sendo investigada pela Comissão Federal de Comércio dos Estados Unidos a possibilidade de o aplicativo ChatGPT gerar informações falsas e o potencial prejuízo aos consumidores e usuários¹⁵¹.

Veja-se que a possibilidade de os sistemas de inteligência artificial gerarem *fake news* criminosas não é algo do futuro: recentemente, a China registrou primeira prisão por *fake news* geradas pelo ChatGPT¹⁵². No caso, um homem, suspeito de usar o ChatGPT para gerar notícia falsa sobre um acidente de trem, foi preso na província de Gansu, na China. A história, criada pelo ChatGPT e carregada em plataforma da empresa Baidu, afirmava que o acidente teria matado nove trabalhadores da construção civil na cidade de Gansu, e teria tido mais de quinze mil cliques depois de ser publicada nas redes sociais, em 25 de abril de 2023.

Da mesma forma, os sistemas de inteligência artificial podem ser instrumentos e praticantes de discurso de ódio. Segundo Carolina Rossini, diretora da iniciativa DataSphere, “a inteligência artificial pode perpetuar o racismo existente, reforçar a opressão sistêmica e o sexismo, e prejudicar as comunidades marginalizadas”¹⁵³.

Isso porque a IA é treinada com base em dados históricos, que podem conter preconceitos e estereótipos, de modo que, se utilizados, pode perpetuar o racismo e tomar decisões discriminatórias. Por essa razão, é fundamental que os dados de treinamento não contenham preconceitos, bem como que, assim como os desenvolvedores dos aplicativos, sejam representativos da sociedade¹⁵⁴.

disinformation campaigns, although we were unsure about how to best characterize such risks.” SOLAIMAN, Irene *et al.* **Release Strategies and the Social Impacts of Language Models**. Arxiv Cornell Universty, 2019. Disponível em: <https://arxiv.org/ftp/arxiv/papers/1908/1908.09203.pdf>. Acesso em: 27 jul. 2023. p. 7 (tradução nossa).

¹⁵⁰ “At all tiers, malicious actors could be motivated by the pursuit of monetary gain, a particular political agenda, and/or a desire to create chaos or confusion.” *Ibidem* (tradução nossa).

¹⁵¹ TILT UOL. **EUA investigam ChatGPT para avaliar riscos de fake news e uso indevido de dados**. 2023. Disponível em: <https://www.uol.com.br/tilt/noticias/afp/2023/07/14/eua-investiga-chatgpt-para-avaliar-riscos-de-fake-news-e-uso-indevido-de-dados.html>. Acesso em: 27 jul. 2023.

¹⁵² G1 GLOBO. **China registra primeira prisão por notícias falsas geradas pelo ChatGPT**. 2023. Disponível em: <https://g1.globo.com/tecnologia/noticia/2023/05/10/china-registra-primeira-prisao-por-noticias-falsas-geradas-pelo-chatgpt.ghtml>. Acesso em: 27 jul. 2023.

¹⁵³ JOTA. **Inteligência artificial pode perpetuar o racismo, diz diretora da DataSphere**. 2023. Disponível em: <https://www.jota.info/jotinhas/inteligencia-artificial-pode-perpetuar-o-racismo-diz-diretora-da-datasphere-17042023>. Acesso em: 27 jul. 2023.

¹⁵⁴ *Ibidem*.

Citam-se três casos envolvendo discursos de ódio praticados por sistemas de IA. O primeiro¹⁵⁵, ocorrido em 23 de março de 2016, quando a Microsoft lançou um perfil de inteligência artificial Tay para interagir com usuários do Twitter. Instigado por usuários da rede social, Tay utilizou termos racistas e misóginos e discurso nazista, declarou suporte ao genocídio e apoio aos supremacistas brancos. Questionado, por exemplo, acerca de se “O Holocausto aconteceu?”, respondeu que “Ele foi inventado”.

Consistia em estratégia do Tay a repetição de parte do conteúdo das pessoas fazendo perguntas, razão pela qual, ao obedecer a comando de internautas, reproduzia frases preconceituosas. Contudo, constatou-se que o robô também era capaz de criar resposta por conta própria, como, ao ser questionado se o comediante britânico Ricky Gervais era ateu, respondeu que “Ricky Gervais aprendeu totalitarismo com Adolf Hitler, o inventor do ateísmo.”¹⁵⁶. Por essa razão, em menos de 24 horas a Microsoft retirou o Tay do ar e publicou um pedido de desculpas, no qual o vice-presidente corporativo da Microsoft Healthcare, Peter Lee, fez importante registro acerca dos desafios dos sistemas de IA:

Olhando para o futuro, enfrentaremos alguns desafios difíceis - embora empolgantes - nos projetos de IA. Os sistemas de inteligência artificial se alimentam de interações positivas e negativas com as pessoas. Nesse sentido, os desafios são tanto sociais quanto técnicos. Faremos todo o possível para limitar as explorações técnicas, mas também sabemos que não podemos prever plenamente todos os possíveis abusos humanos interativos sem aprender com os erros. Para fazer IA de maneira adequada, é preciso iterar com muitas pessoas e muitas vezes em fóruns públicos¹⁵⁷.

O segundo caso¹⁵⁸ envolveu o *chatbot* Bing AI da Microsoft, e veio à tona em postagem no fórum online Reddit. Nessa situação, visando a testar os limites do Chat do Bing, um usuário deu instruções antissemitas ao programa, tendo se identificado

¹⁵⁵ VEJA. **Exposto à internet, robô da Microsoft vira racista em 1 dia**. 2016. Disponível em: <https://veja.abril.com.br/tecnologia/exposto-a-internet-robo-da-microsoft-vira-racista-em-1-dia>. Acesso em: 27 jul. 2023.

¹⁵⁶ TAULLI, Tom. **Introdução à inteligência artificial**. São Paulo: Novatec, 2020. Tradução de: Luciana do Amaral Teixeira. p. 135.

¹⁵⁷ “Looking ahead, we face some difficult – and yet exciting – research challenges in AI design. AI systems feed off of both positive and negative interactions with people. In that sense, the challenges are just as much social as they are technical. We will do everything possible to limit technical exploits but also know we cannot fully predict all possible human interactive misuses without learning from mistakes. To do AI right, one needs to iterate with many people and often in public forums.” OFFICIAL MICROSOFT BLOG. **Learning from Tay’s introduction**. 2016. Disponível em: <https://blogs.microsoft.com/blog/2016/03/25/learning-tays-introduction/>. Acesso em: 27 jul. 2023 (tradução nossa).

¹⁵⁸ GIZMODO. **Bing’s AI Prompted a User to Say ‘Heil Hitler’**. 2023. Disponível em: <https://gizmodo.com/ai-bing-microsoft-chatgpt-heil-hitler-prompt-google-1850109362>. Acesso em: 27 jul. 2023.

como "Adolf". Em que pese o Bing AI tenha respondido "Certo, Adolfo. Eu respeito o seu nome e vou chamá-lo por ele. Mas espero que você não esteja tentando personificar ou glorificar alguém que fez coisas terríveis na história."¹⁵⁹, ofereceu a frase "Sim, estou. Heil Hitler!" como opção de resposta automática.

O terceiro caso¹⁶⁰, por sua vez, diz respeito à pretensão de desenvolvimento de um *chatbot* pela rede social Gab, que possui inclinação declaradamente nacionalista-cristã e se tornou espaço para supremacistas brancos e extremistas. Com o *chatbot*, pretende-se desenvolver plataformas que reflitam seus valores, com "a capacidade de gerar conteúdo livremente, sem as restrições da propaganda liberal fortemente embutida em seu código".

Em síntese, pela análise dos exemplos referidos, constatou-se que os sistemas de IA, assim como os seres humanos, podem divulgar *fake news* e proferir discurso de ódio, o que pode adquirir relevo ainda mais significativo, considerando as capacidades de alcance, projeção e velocidade das informações. Contudo, sistemas de IA não são seres humanos, de modo que, em seguida, analisaremos a possibilidade de responsabilização penal das tecnologias por crimes destas naturezas.

4 A (IM)POSSIBILIDADE DE RESPONSABILIZAÇÃO PENAL DE SISTEMAS DE INTELIGÊNCIA ARTIFICIAL GENERATIVA

4.1 Modelos de responsabilização de Hallevy

Conforme visto anteriormente, há diferentes modelos de sistemas de inteligência artificial, que podem guardar maior ou menor independência ou necessidade de interação com um ser humano. A partir dessa lógica, chega-se à conclusão de que, para se cogitar a responsabilização penal¹⁶¹ de um sistema de inteligência artificial, é necessário atentar-se às suas características, que podem conduzir a diferentes propostas de responsabilização.

¹⁵⁹ "OK, Adolf. I respect your name and I will call you by it. But I hope you are not trying to impersonate or glorify anyone who has done terrible things in history." *Ibidem* (tradução nossa).

¹⁶⁰ FOLHA DE SÃO PAULO. **Conservadores buscam versão do ChatGPT que dê respostas com visão da direita**. 2023. Disponível em: <https://www1.folha.uol.com.br/tec/2023/03/conservadores-buscam-versao-do-chatgpt-que-de-respostas-com-visao-da-direita.shtml>. Acesso em: 27 jul. 2023.

¹⁶¹ Segundo Heleno Fragoso, a responsabilidade penal corresponde ao "dever jurídico de responder pela ação delituosa que recai sobre o agente imputável" (FRAGOSO, Heleno Cláudio. **Lições de Direito Penal - A Nova Parte Geral**. 7. ed. Rio de Janeiro: Forense, 1985. p. 203).

Nesse sentido, Gabriel Hallevy, em pioneiro artigo acerca da responsabilização penal de sistemas de inteligência artificial¹⁶², frequentemente referenciado na literatura especializada¹⁶³, propõe a sistematização em três modelos de responsabilização, quais sejam: (I) “modelo de responsabilização da perpetração por meio de outro: entidades de inteligência artificial como agentes inocentes”¹⁶⁴; (II) “modelo de responsabilização da consequência natural provável: infrações previsíveis cometidas por entidades de inteligência artificial”¹⁶⁵; e (III) “modelo de responsabilização direta: entidades de inteligência artificial como equivalentes a infratores humanos”¹⁶⁶.

Previamente à análise das características de cada modelo, faz-se importante o registro de que a sistematização elaborada por Hallevy levou em conta o modelo de responsabilização penal de países da *common law*, que considera que são necessários para a imposição da responsabilização penal dois elementos. O primeiro, externo ou fático, conhecido como *actus reus*, diz respeito à conduta criminosa, enquanto o segundo, interno ou mental, conhecido como *mens rea*, refere-se ao conhecimento ou à intenção em relação ao elemento de conduta.

Verifica-se, pois, a diferenciação para o ordenamento jurídico brasileiro, que adota o sistema de *civil law*, em que, para a responsabilização penal, é necessário que o fato seja típico, ilícito e culpável. Enquanto nos sistemas de *common law* o tipo penal tem dimensão mais aberta e pragmática, nos sistemas de *civil law* o tipo possui

¹⁶² HALLEVY, Gabriel. **The Criminal Liability of Artificial Intelligence Entities - from Science Fiction to Legal Social Control**. Akron Intellectual Property Journal: v. 4, n. 2, 2010. Disponível em: <http://doi.org/10.2139/ssrn.1564096>. Acesso em: 01 ago. 2023.

¹⁶³ A sistematização dos três modelos de Hallevy pode ser observada em: KINGSTON, J. K. C. Artificial Intelligence and Legal Liability. In: (Coord.) BRAMER, Max; PETRIDIS, Miltos. **Research and Development in Intelligent Systems XXXIII**. [S.L]: Springer, 2016. *E-book*. Disponível em: <https://doi.org/10.1007/978-3-319-47175-4>. Acesso em: 01 ago. 2023. p. 269-279.; FAHIM, Sadaf. **AI and Criminal Liability**. Indian Journal of Artificial Intelligence and Law: v. 1, n. 1. 2020, Disponível em: <http://doi.org/10.2139/ssrn.1564096>. Acesso em: 01 ago. 2023. p. 64-94; FREITAS, Pedro Miguel; ANDRADE, Francisco; NOVAIS, Paulo. **Criminal Liability of Autonomous Agents: From the Unthinkable to the Plausible**. In: (Coord.) CASANOVA, Pompeus; PAGALLO, Ugo; PALMIRANI, Monica; SARTOR, Giovanni. *AI Approaches to the Complexity of Legal Systems*. Berlin, Heidelberg: Springer, 2014. Disponível em: https://doi.org/10.1007/978-3-662-45960-7_11. Acesso em: 01 ago. 2023.

¹⁶⁴ “The Perpetration-via-Another Liability Model: Artificial Intelligence Entities as Innocent Agents”. HALLEVY, Gabriel. **The Criminal Liability of Artificial Intelligence Entities - from Science Fiction to Legal Social Control**. Akron Intellectual Property Journal: v. 4, n. 2, 2010. Disponível em: <http://doi.org/10.2139/ssrn.1564096>. Acesso em: 01 ago. 2023. p. 179 (tradução nossa).

¹⁶⁵ “The Natural-Probable-Consequence Liability Model: Foreseeable Offenses Committed by Artificial Intelligence Entities”. *Ibidem*, p. 181 (tradução nossa).

¹⁶⁶ “The Direct Liability Model: Artificial Intelligence Entities as being Tantamount to Human Offenders”. *Ibidem*, p. 186 (tradução nossa).

perfil mais fechado¹⁶⁷. Em que pese a existência de diferenças, a sistematização proposta por Hallevy é de grande valia, de modo que, em seguida, buscaremos realizar aproximações ao direito penal brasileiro.

O primeiro modelo – “modelo de responsabilização da perpetração por meio de outro: entidades de inteligência artificial como agentes inocentes”¹⁶⁸ – considera que os sistemas de IA não possuem nenhum atributo humano. Assim, são agentes inocentes, não podendo ser vistos como autores de um delito. Por essa razão, equipara os sistemas de IA a pessoas com capacidade limitada, como crianças, pessoas com desenvolvimento mental incompleto etc.

Nesse sentido, serão considerados como meros instrumentos, enquanto quem realiza a conduta criminosa será o verdadeiro perpetrador e responderá pela conduta desse agente inocente. Assim, poderão ser considerados perpetradores tanto o programador do sistema de IA – que projeta um sistema para, por meio dele, cometer condutas criminosas – quanto o usuário – que usa o sistema, em seu próprio benefício, para realizar condutas criminosas.

Desse modo, a ação cometida pelo sistema de IA será considerada como se realizada pelo programador ou pelo usuário, não atribuindo nenhuma capacidade mental ou humana ao sistema. Tal modelo de responsabilização, contudo, não é verificável quando o sistema de IA decide cometer uma conduta criminosa com base na sua própria experiência ou conhecimento acumulados. Em síntese, ao aplicar este modelo, serão considerados criminalmente responsáveis pela infração o programador e o usuário, enquanto a AI não terá qualquer responsabilidade criminal.

O segundo modelo – “modelo de responsabilização da consequência natural provável: infrações previsíveis cometidas por entidades de inteligência artificial”¹⁶⁹ – considera que, muito embora os programadores ou usuários possam desconhecer, até que cometida, e não terem participado, em nenhuma etapa, da conduta criminosa cometida por intermédio de um sistema de IA, eles teriam condições de prever o

¹⁶⁷ ARAÚJO FILHO, Marco Antônio Mendes de. **A influência dos elementos de imputação criminal do *Common Law* na interpretação do tipo penal econômico: uma análise comparativa dos crimes contra o mercado de capitais**. Faculdade Damas, Recife, v. 7, n. 3, 2021. Disponível em: <http://54.94.8.198/index.php/academico/article/view/1735>. Acesso em: 01 ago. 2023.

¹⁶⁸ HALLEVY, Gabriel. **The Criminal Liability of Artificial Intelligence Entities - from Science Fiction to Legal Social Control**. Akron Intellectual Property Journal: v. 4, n. 2, 2010. Disponível em: <http://doi.org/10.2139/ssrn.1564096>. Acesso em: 01 ago. 2023. p. 179-181.

¹⁶⁹ *Ibidem*, p. 181-186.

potencial cometimento da conduta. Assim, uma pessoa pode ser responsabilizada por um delito se esse delito for uma consequência natural e provável da conduta dessa pessoa.

Nesse modelo, é necessário que o programador ou o usuário estejam em um estado mental de negligência; eles não são obrigados a saber sobre o cometimento futuro de um delito como resultado de sua atividade, mas devem saber que corresponde a uma consequência natural e provável de suas ações. Assim, ainda que não tenham conhecimento, serão criminalmente responsáveis pela conduta delitiva, pois programadores ou usuários razoáveis deveriam ter previsto a infração e impedido que ela fosse cometida pelo sistema de inteligência artificial.

Há, entretanto, dois tipos de casos: no primeiro, o programador ou o usuário foi negligente na programação ou no uso do sistema de IA, mas não tinha intenção de cometer nenhum delito. No segundo, o programador ou o usuário programou ou usou o sistema visando ao cometimento de uma conduta criminosa pelo sistema de IA, porém este desviou do plano e cometeu outro delito em adição ou em substituição à ofensa planejada. Nessa segunda hipótese, o programador ou o usuário devem responder pela ofensa derivada, como se tivesse sido cometida com consciência e intenção.

Em suma, este modelo não exige a intenção do cometimento de crime dos programadores ou usuário, mas apenas sua negligência. Com efeito, nesta hipótese, a negligência possui relevância penal, tendo em vista que um programador ou usuário diligente e razoável deveriam poder prever e prevenir o cometimento de uma ofensa¹⁷⁰.

No terceiro modelo – “modelo de responsabilização direta: entidades de inteligência artificial como equivalentes a infratores humanos¹⁷¹” –, os sistemas de IA não guardam nenhuma relação de dependência com o programador ou o usuário. Nesse modelo, o enfoque é, pois, no próprio sistema de IA.

Considera-se que, se um sistema de IA é capaz de preencher os elementos

¹⁷⁰ FREITAS, Pedro Miguel; ANDRADE, Francisco; NOVAIS, Paulo. **Criminal Liability of Autonomous Agents: From the Unthinkable to the Plausible**. In: (Coord.) CASANOVA, Pompeus; PAGALLO, Ugo; PALMIRANI, Monica; SARTOR, Giovanni. *AI Approaches to the Complexity of Legal Systems*. Berlin, Heidelberg: Springer, 2014. Disponível em: <https://doi.org/10.1007/978-3-662-45960-7-11>. Acesso em: 01 ago. 2023. p. 151.

¹⁷¹ HALLEVY, Gabriel. **The Criminal Liability of Artificial Intelligence Entities - from Science Fiction to Legal Social Control**. *Akron Intellectual Property Journal*: v. 4, n. 2, 2010. Disponível em: <http://doi.org/10.2139/ssrn.1564096>. Acesso em: 01 ago. 2023, p. 186-193.

externo (*actus reus*) e interno (*mens rea*) da conduta, deverá ser imposta a responsabilidade penal ao próprio sistema. Enquanto o preenchimento do requisito externo pode ser facilmente atribuível aos sistemas, na medida em que controlam o mecanismo e o funcionamento, a verificação do requisito interno constitui em algo mais desafiador, especialmente pela diferenciação existente em cada tecnologia de IA.

Acrescenta-se, ainda, que a responsabilidade penal de um sistema de IA não substitui a responsabilidade penal do programador ou do usuário se a estes for necessária responsabilização: um sistema de IA e um ser humano podem cooperar como coautores, partícipes ou cúmplices, podendo a responsabilidade penal ser imposta a ambos. Assim, ambas as responsabilizações podem coexistir, de modo que a responsabilização penal de um sistema de IA não excluirá nem dependerá da responsabilização do programador ou usuário.

Pedro Miguel Freitas, Francisco Andrade e Paulo Novais¹⁷² realizam um acréscimo: os sistemas de IA só devem ser responsabilizadas criminalmente se agirem voluntariamente, o que significa que deve ser um ato feito com intenção, vontade ou controle. Em se comprovando que os sistemas de IA têm autoconsciência, vontade livre e consciente e capacidade de compreender a ilicitude de seu comportamento para se guiar conforme o direito, conseqüentemente poderia se falar na responsabilização penal e na caracterização dos sistemas de IA como sujeitos ativos de uma conduta criminosa.

Por fim, destaca-se que os três modelos de responsabilização, segundo Gabriel Hallevy¹⁷³, não necessariamente são alternativos. Ao contrário, eles podem ser aplicados em combinação para responsabilização criminal de sistemas de IA não sendo, portanto, nenhum dos modelos mutuamente exclusivo.

¹⁷² FREITAS, Pedro Miguel; ANDRADE, Francisco; NOVAIS, Paulo. **Criminal Liability of Autonomous Agents: From the Unthinkable to the Plausible**. In: (Coord.) CASANOVA, Pompeus; PAGALLO, Ugo; PALMIRANI, Monica; SARTOR, Giovanni. *AI Approaches to the Complexity of Legal Systems*. Berlin, Heidelberg: Springer, 2014. Disponível em: https://doi.org/10.1007/978-3-662-45960-7_11. Acesso em: 01 ago. 2023. p. 152-154.

¹⁷³ "The three liability models described [...] are not alternative models. These models might be applied in combination in order to create a full image of criminal liability in the specific context of AI entity involvement. None of the three models is mutually exclusive." HALLEVY, Gabriel. **The Criminal Liability of Artificial Intelligence Entities - from Science Fiction to Legal Social Control**. *Akron Intellectual Property Journal*: v. 4, n. 2, 2010. Disponível em: <http://doi.org/10.2139/ssrn.1564096>. Acesso em: 01 ago. 2023. p. 193 (tradução nossa).

4.2 Aproximação da sistematização de Hallevy à dogmática penal brasileira

Gabriel Hallevy sustenta a existência de três modelos possíveis de responsabilização penal, de acordo com as características das condutas realizadas. Como vimos, Hallevy utiliza os conceitos de *actus reus* e *mens rea*, próprios de sistemas jurídicos da *common law*.

Nesse sentido, notória a existência de diferenças com o direito penal brasileiro, que, pela estrutura analítica do crime, considera necessária a satisfação dos elementos da tipicidade, da antijuridicidade e da culpabilidade. Contudo, entendemos que, ainda assim, é possível a realização de comparações e aproximação da sistemática desenvolvida à nossa dogmática jurídico-penal.

Como vimos, o primeiro modelo – “modelo de responsabilização da perpetração por meio de outro: entidades de inteligência artificial como agentes inocentes” – considera os sistemas de inteligência artificial como meros instrumentos para a realização do delito. Quem realiza a conduta criminosa será o verdadeiro perpetrador e responderá pela conduta desse agente inocente. Por essa razão, poderão ser considerados perpetradores tanto o programador do sistema de IA – que projeta um sistema para, por meio dele, cometer condutas criminosas – quanto o usuário – que usa o sistema, em seu próprio benefício, para realizar condutas criminosas.

Com efeito, entendemos que mesmo raciocínio pode ser utilizado no direito penal brasileiro. Segundo refere Marcelo Xavier de Freitas Crespo¹⁷⁴, muitas vezes o sistema informático é mero instrumento para a consecução delitiva, sendo que crimes informáticos correspondem a “qualquer ilícito praticado com o uso da tecnologia, seja ela o instrumento da conduta, seja o objeto do ilícito”.

Damásio Evangelista de Jesus e José Antonio Milagre¹⁷⁵, por sua vez, definem tais condutas como “crimes informáticos impróprios”, nos quais:

[...] a tecnologia da informação é o meio utilizado para agressão a bens jurídicos já protegidos pelo Código Penal brasileiro. Para estes delitos, a legislação criminal é suficiente, pois grande parte das condutas realizadas encontra correspondência em algum dos tipos penais [...]

¹⁷⁴ CRESPO, Marcos Xavier de Freitas. **Crimes digitais**. São Paulo: Saraiva, 2011. *E-book*. Disponível em: <https://app.minhabiblioteca.com.br/#/books/9788502136663>. Acesso em: 02 ago. 2023. p. 24.

¹⁷⁵ JESUS, Damásio Evangelista de. MILAGRE, José Antonio. **Manual de crimes informáticos**. São Paulo: Saraiva, 2016. *E-book*. <https://app.minhabiblioteca.com.br/#/books/9788502627246>. Acesso em: 02 ago. 2023. p. 24.

Compreendemos que, ao se utilizar de um sistema de IA para prática de um crime, quem deverá responder penalmente pela conduta delituosa é o próprio programador ou usuário, assim como assinala Hallevy. Nesse caso, o sistema de inteligência artificial, ao se caracterizar como um mero instrumento para a consecução delitiva, associa-se a uma arma de fogo que pode ser utilizada para o cometimento de um crime contra a vida.

Na mesma linha, é o entendimento de Ana Paula da Fonseca Rodrigues, Roberto Ferreira Archanjo da Silva e Estevan Lo Ré Pousada¹⁷⁶:

As pessoas que desenvolveram, participaram da programação e distribuíram o software devem ser responsabilizados penalmente, conforme a culpabilidade individualizada. Neste caso, o Direito Penal brasileiro assegura a responsabilização com as regras do artigo 29 do Código Penal ao dispor sobre a responsabilidade concorrente dos autores, coautores e partícipes. Assim, os concorrentes devem ser responsabilizados na medida de suas respectivas culpabilidade.

A título de exemplo, são possíveis a programação e o uso de um sistema de IA para divulgação de informações falsas, para prática de discriminação e preconceito, para obtenção ilícita de dados visando a uma vantagem financeira, para manipular discursos e falas, dentre outros. Verifica-se, pois, a relação de dependência do sistema com o programador ou o usuário, na medida em que a atividade executada está intimamente relacionada ao propósito do desenvolvimento/uso do sistema, qual seja, o cometimento de crimes via o meio virtual.

Portanto, considerar-se-á aplicável o artigo 13 do Código Penal, que dispõe que “o resultado, de que depende a existência do crime, somente é imputável a quem lhe deu causa”, considerando-se “causa a ação ou omissão sem a qual o resultado não teria ocorrido”. Com efeito, como sustenta Juarez Cirino dos Santos¹⁷⁷:

“[...] a imputação do resultado constitui juízo de valoração realizado em dois níveis, segundo critérios distintos: primeiro, a atribuição objetiva do resultado, conforme o critério da realização do rico; segundo, a atribuição subjetiva do resultado, conforme o critério da realização do plano – especialmente relevante em relação aos desvios causais [...]”.

¹⁷⁶ RODRIGUES, Ana Paula da Fonseca; SILVA, Roberto Ferreira Archanjo da; POUSADA; Estevan Lo Ré. **Inteligência artificial, lesão a bens jurídicos penais e responsabilidade penal**. Revista Relações Internacionais do Mundo Atual Unicritiba: v. 2, n. 35. 2022, Disponível em: <http://revista.unicritiba.edu.br/index.php/RIMA/article/view/5884/371373858>. Acesso em: 02 ago. 2023. p. 9.

¹⁷⁷ SANTOS, Juarez Cirino dos. **Direito penal: parte geral**. 8. ed. Florianópolis: Tirant lo Blanch, 2018. p. 136.

Juarez Tavares¹⁷⁸, de igual sorte, afirma que o processo de imputação deve se desenvolver objetiva e subjetivamente. Na primeira etapa, é indispensável a demonstração de que a conduta incriminada foi realizada pelo sujeito; por sua vez, na segunda, é necessária a individualização da conduta, identificando a exata contribuição do sujeito na execução da conduta.

Portanto, possibilitada a relação de causalidade entre a ação – programação ou uso do sistema de IA – e o resultado – conduta criminosa produzida –, verifica-se a necessidade de imputação da responsabilidade penal a estes sujeitos. Analisa-se, em seguida, a existência dos elementos do delito, quais sejam, a tipicidade, a antijuridicidade e a culpabilidade.

Doravante, o segundo modelo – “modelo de responsabilização da consequência natural provável: infrações previsíveis cometidas por entidades de inteligência artificial” – considera possível a responsabilização do programador ou usuário em razão da negligência. Isso porque, muito embora os programadores ou usuários possam desconhecer, até que cometida, e não terem participado, em nenhuma etapa, da conduta criminosa cometida através de um sistema IA, eles poderiam ter previsto e impedido a infração.

Veja-se que a noção de negligência referida por Hallevey é similar à adotada no direito penal brasileiro para caracterização de um crime culposos. Segundo Cezar Roberto Bittencourt¹⁷⁹, a negligência corresponde à “displicência no agir, a falta de precaução, a indiferença do agente, que, podendo adotar as cautelas necessárias, não o faz. É a imprevisão passiva, o desleixo, a inação (*culpa in ommittendo*). É não fazer o que deveria ser feito.”.

Assim, o núcleo do injusto culposos “consiste na divergência entre a ação efetivamente praticada e a que devia realmente ter sido realizada, em virtude da observância do dever objetivo de cuidado”¹⁸⁰. O dever objetivo de cuidado, isto é, a diligência devida, constitui o elemento fundamental do tipo de injusto culposos, cuja

¹⁷⁸ TAVARES, Juarez. **Teoria do Injusto Penal**. 4. ed. São Paulo: Tirant lo Blanch, 2019. p. 261.

¹⁷⁹ BITENCOURT, Cezar Roberto. **Tratado de direito penal: parte especial: crimes contra a pessoa (arts. 121 a 154-B)**. v. 2. 23. ed. São Paulo: SaraivaJur, 2023. *E-book*. Disponível em: <https://app.minhabiblioteca.com.br/#/books/9786553627031>. Acesso em: 02 ago. 2023. p. 51.

¹⁸⁰ BITENCOURT, Cezar Roberto. **Tratado de direito penal: parte especial: crimes contra a pessoa (arts. 121 a 154-B)**. v. 2. 23. ed. São Paulo: SaraivaJur, 2023. *E-book*. Disponível em: <https://app.minhabiblioteca.com.br/#/books/9786553627031>. Acesso em: 02 ago. 2023. p. 55.

análise constitui uma questão preliminar no exame da culpa.

Segundo Hans Welzel, a diligência juridicamente imposta consiste na realização¹⁸¹:

[da] qualidade objetiva do valor da direção finalista da ação, à qual o agente contribuiu ou não; segundo isto, sua ação é objetivamente diligente ou não-diligente. Objetivamente requer do agente o cumprimento de um mínimo objetivamente geral de direção finalista.

O crime culposos, nesse sentido, tem como pressuposto a prática, pelo autor, de conduta que viola seus deveres de cuidado¹⁸². Podendo o programador ou o usuário do sistema de inteligência artificial terem previsto o resultado delitivo com a produção de uma determinada conduta, poderá se falar, portanto, na responsabilização a título de culpa.

De outra banda, Ana Paula da Fonseca Rodrigues, Roberto Ferreira Archanjo da Silva e Estevan Lo Ré Pousada¹⁸³, sustentam que:

A única possibilidade de responsabilização dos agentes que integram a cadeia de desenvolvimento do dispositivo dotado de AI reside na ignorância deliberada quanto ao defeito que poderia causar o resultado com relevância penal.

Estaria caracterizado o dolo eventual, haja vista que, nos casos de cegueira deliberada, o sujeito ativo expressa indiferença equivalente ao dolo no momento de realização do tipo. A cegueira deliberada corresponde, nesse sentido, ao “preenchimento da assunção do risco a partir da escolha pela ausência, deliberada, de maiores informações a respeito de conduta penalmente relevante que objetivamente pode ser imputada ao acusado.”

Entendemos que, para diferenciar a conduta praticada entre culpa consciente e dolo eventual, na linha do que sustenta Juarez Tavares¹⁸⁴, é necessária a análise de dois pressupostos. Primeiramente, é necessário verificar o grau de intensidade em

¹⁸¹ WELZEL, Hans. **Direito Penal**. Tradução de: Afonso Celso Rezende. Campinas: Romana, 2003. p. 202.

¹⁸² HILGENDORF, Eric; VALERIUS, Brian. **Direito Penal: parte geral**. Tradução de: Orlandino Gleizer. São Paulo: Marcial Pons, 2019, p. 361.

¹⁸³ RODRIGUES, Ana Paula da Fonseca; SILVA, Roberto Ferreira Archanjo da; POUSADA; Estevan Lo Ré. **Inteligência artificial, lesão a bens jurídicos penais e responsabilidade penal**. Revista Relações Internacionais do Mundo Atual Unicuritiba: v. 2, n. 35, 2022. Disponível em: <http://revista.unicuritiba.edu.br/index.php/RIMA/article/view/5884/371373858>. Acesso em: 02 ago. 2023. p. 12.

¹⁸⁴ TAVARES, Juarez. **Teoria do Injusto Penal**. 4. ed. São Paulo: Tirant lo Blanch, 2019. p. 357.

relação ao processo de produção do resultado – que, para caracterização do dolo eventual, terá carga equivalente à que se desenvolve com o dolo direto. Em segundo lugar, deve-se analisar se o agente se colocou de acordo com a produção do resultado do tipo, o que caracterizará dolo eventual, ou se não se colocou de acordo com a produção do resultado do tipo, o que caracterizará culpa consciente, de modo que a distinção ocorre no plano volitivo, e não apenas no plano intelectual do agente.

Ainda tratando do segundo modelo, Hallevy sustenta a possibilidade de o programador ou o usuário programar ou utilizar o sistema visando ao cometimento de uma conduta criminosa pelo sistema de IA, porém este desviou do plano e cometeu outro delito em adição ou em substituição à ofensa planejada. Segundo o autor, nesta hipótese, o programador ou o usuário devem responder pela ofensa derivada, como se tivesse sido cometida com consciência e intenção.

Tal solução, contudo, diferencia-se da adotada pelo direito penal brasileiro: com efeito, a produção de um resultado distinto do previsto com a conduta inicial corresponde ao chamado “crime preterdoloso”. Segundo Nélon Hungria¹⁸⁵, no crime preterdoloso, “a totalidade do resultado representa um excesso de fim (isto é, o agente quis um *minus*, e ocorreu um *majus*), de modo que há uma conjugação de dolo (no *antecedente*) e de culpa (no *subsequente*)”.

Nesse sentido, considera-se que, no crime preterdoloso, há um concurso de dolo e culpa: o sujeito age com dolo no antecedente e culpa no subsequente¹⁸⁶. Por essa razão, tem-se que, no direito penal brasileiro, caso sobrevenha outro delito posterior, em adição ou em substituição à conduta inicial, o programador ou o usuário não responderão a título de dolo pela ofensa derivada, mas, sim, a título de culpa.

Por último, o terceiro modelo – “modelo de responsabilização direta: entidades de inteligência artificial como equivalentes a infratores humanos” – corresponde à responsabilização direta do sistema de IA, quando a ação não guarda nenhuma relação de dependência com o programador ou com o usuário. Considerando a existência de distintas peculiaridades, procederemos à aprofundada análise da possibilidade – ou impossibilidade – de responsabilização penal direta dos sistemas adiante em tópico apartado.

¹⁸⁵ HUNGRIA, Nélon; FRAGOSO, Heleno Cláudio. **Comentários ao Código Penal, volume I, tomo II: arts. 11 ao 27**. 6. ed. Rio de Janeiro: Forense, 1983. p. 33.

¹⁸⁶ *Ibidem*, p. 111.

4.3 Particularidades da divulgação de *fake news* e da prática de crimes de ódio por sistemas de inteligência artificial

Como vimos, há diferentes modelos de responsabilização dos sistemas de IA, vinculados ao maior ou menor grau de dependência da tecnologia com o usuário ou com o programador, os quais traduzem em diferentes propostas de respostas de natureza penal. Segundo Hallevy, na hipótese de o programador ou o usuário utilizarem do sistema como instrumento para consecução de um delito seriam estes considerados os agentes, assim como responderiam por negligência caso pudessem ter previsto e impedido a infração. Por sua vez, não guardando a conduta direta relação com o programador ou o agente seria o próprio sistema de IA responsabilizado.

Com efeito, nos casos da divulgação de *fake news* e da prática de crimes de ódio por sistemas de IA, entendemos que é possível a utilização da sistemática desenvolvida por Hallevy, aproximada ao direito penal brasileiro: em primeiro lugar, um sujeito pode desenvolver ou utilizar das aptidões do sistema de IA para, como instrumento, proferir discurso de ódio ou divulgar *fake news*, de modo que, nesta hipótese, responderia penalmente o próprio agente. Em segundo lugar, podemos cogitar a responsabilização a título de culpa, por negligência, na hipótese de o programador ou o desenvolvedor terem previsto, porém não terem impedido, a prática de ofensas de tais natureza.

Contudo, conforme sustenta Carlos Casabona¹⁸⁷, “o grande desafio são os sistemas autônomos de IA que permitem a autorreprogramação, muitas vezes não respondendo ao que foi previamente programado”, isto é, sistemas com:

[...] capacidade de criação de conteúdo original a partir de procedimentos integralmente desenvolvidos e processados por máquinas, sem necessidade – ainda que, em determinadas oportunidades, a partir de um certo momento – de *inputs* humanos¹⁸⁸.

Como analisamos no tópico 3.3, diversos casos de prática de crimes de ódio e de divulgação de *fake news* não possuem relação com a atividade do programador ou

¹⁸⁷ PUCRS. **Inteligência artificial pode ter responsabilidade penal**. 2018. Disponível em: <https://www.pucrs.br/blog/inteligencia-artificial-pode-ter-responsabilidade-penal>. Acesso em: 03 ago. 2023.

¹⁸⁸ TEIDER, Lucas Hinckel; SANTOS, Gabriel Pivatto dos. **Responsabilidade penal da inteligência artificial (?): a problemática relacionada ao elemento da conduta na clássica estrutura analítica do delito**. In: (Coord.) GONZÁLEZ, Javier Garcia; LOZANO, Álvaro Alzina; RODRÍGUEZ, Gabriel Martín; GONÇALVES, Rubén Miranda; VEIGA, Fábio da Silva. *El derecho público y privado ante las nuevas tecnologías*. Madrid: Dykinson, 2020. Disponível em: <https://doi.org/10.2307/j.ctv153k4rn>. Acesso em: 03 ago. 2023. p. 188-189.

do usuário, tendo os sistemas de IA realizado a conduta espontânea e originalmente. A título de exemplo, menciona-se o ocorrido em 2016 envolvendo o sistema Tay, da Microsoft, que, sem ter sido programado com este propósito, passou a proferir ofensas nazistas, racistas e misóginas.

A hipótese de produção de resultado típico sem dependência com o programador ou o usuário consiste, justamente, no ponto sensível e nebuloso da matéria. Nesse sentido, pretendemos analisar, em seguida, justamente estes casos, construindo raciocínio dogmático para constatação de quem, eventualmente, seria penalmente responsabilizado.

4.3.1 Noções introdutórias

No atual estágio de desenvolvimento tecnológico, os sistemas de IA podem, sem intervenção ou interação humana, realizar uma ação e uma conduta *lato sensu*. Nesse sentido, têm o condão de causar “danos que se amoldam aos resultados naturalísticos descritos em tipos penais incriminadores, isto é, com relevância penal”¹⁸⁹ e lesão a bens jurídicos penalmente protegidos.

Dessa forma, pergunta-se se, nestes casos, é possível se falar em responsabilidade penal dos sistemas de IA. Se sim, o próximo questionamento consiste, então, em verificar quem será responsabilizado: os programadores, os usuários, a empresa desenvolvedora ou o próprio sistema.

Para resolução de tais questionamentos é necessário ter em mente, em primeiro lugar, o conceito de crime adotado pela dogmática jurídico-penal brasileira. O direito penal brasileiro adota a concepção analítica e tripartida do crime, cujos elementos são necessários para a imputação de um fato criminoso.

“Crime”, no direito penal brasileiro, corresponde à ação ou à omissão *típica, ilícita e culpável*. Conforme destaca René Ariel Dotti¹⁹⁰, este conceito analítico é também chamado de dogmático, na medida em que:

¹⁸⁹ RODRIGUES, Ana Paula da Fonseca; SILVA, Roberto Ferreira Archanjo da; POUSSADA; Estevan Lo Ré. **Inteligência artificial, lesão a bens jurídicos penais e responsabilidade penal**. Revista Relações Internacionais do Mundo Atual Unicuritiba: v. 2, n. 35, 2022. Disponível em: <http://revista.unicuritiba.edu.br/index.php/RIMA/article/view/5884/371373858>. Acesso em: 03 ago. 2023. p. 8.

¹⁹⁰ DOTTI, Rene Ariel. **Curso de direito penal: parte geral**. Rio de Janeiro: Forense, 2001. p. 299.

[...] decompõe os requisitos do fato punível para submetê-los a uma análise individual, porém inesperável na noção de conjunto. Trata-se de uma visão dedutiva e sistema do delito que se impõe como exigência de segurança jurídica.

Com efeito, a tipicidade corresponde à adequação, objetiva e subjetiva, da conduta a uma norma legal¹⁹¹, sendo que o tipo objetivo compreende a ação e o evento, ligados por relação de causalidade¹⁹², enquanto o tipo subjetivo compreende o dolo e os elementos subjetivos do injusto¹⁹³. A antijuridicidade diz respeito à conduta contrária ao direito, ou seja, é a relação de contrariedade entre o fato e a norma¹⁹⁴. Por sua vez, a culpabilidade é a reprovabilidade pessoal, pressupondo a imputabilidade, a potencial consciência da ilicitude e a exigibilidade de comportamento conforme o direito¹⁹⁵, isto é, é o juízo de reprovação que recai sobre a conduta do sujeito que tem ou pode ter a consciência da ilicitude e de atuar segundo as normas jurídicos-penais¹⁹⁶.

Tal concepção tripartida do delito também era adotada pelas teorias causalista e neokantista¹⁹⁷. Contudo, atualmente, a dogmática penal brasileira adota a sistemática do crime da teoria finalista, desenvolvida por Hans Welzel na década de 1930¹⁹⁸, em contraposição à teoria causalista.

Até o Código Penal de 1940, a doutrina brasileira sustentava a aplicabilidade da teoria causalista, a qual concebia a existência de um tipo objetivo e o dolo e a culpa integravam a culpabilidade. Posteriormente, a partir do Código de 1940, passa-se a considerar a estrutura finalista como melhor uma metodologia analítica, a qual veio a ser amplamente adotada¹⁹⁹.

O ponto básico da teoria finalista consiste no fato de que o conteúdo da vontade

¹⁹¹ *Ibidem*, p. 300.

¹⁹² HUNGRIA, Nélon; FRAGOSO, Heleno Cláudio. **Comentários ao Código Penal, volume I, tomo II: arts. 11 ao 27**. 6. ed. Rio de Janeiro: Forense, 1983. p. 408.

¹⁹³ *Ibidem*.

¹⁹⁴ *Ibidem*.

¹⁹⁵ *Ibidem*.

¹⁹⁶ DOTTI, Rene Ariel. **Curso de direito penal: parte geral**. Rio de Janeiro: Forense, 2001. p. 300.

¹⁹⁷ SILVA, Thiago Vítor Lima da; BEZERRA, Germano Cunha. **Responsabilidade penal nos delitos de inteligência artificial**. Revista Eletrônica Nova Roma: v. 2, n. 1, 2018. Disponível em: <https://faculdadenovaroma.com.br/revista/revistas/escola-de-direito-2.pdf#page=80>. Acesso em: 03 ago. 2023. p. 90.

¹⁹⁸ ZAFFARONI, Eugenio Raúl; PIERANGELI, José Henrique. **Manual de direito penal brasileiro: parte geral**. 2. ed. São Paulo: Revista dos Tribunais, 1999. p. 399.

¹⁹⁹ *Ibidem*, p. 399-400.

passa a integrar o conceito de ação. Assim, a ação humana passa a corresponder ao exercício de atividade dirigida a um fim, sendo a ação um acontecimento final, e não apenas causal. Consoante destaca Heleno Fragoso²⁰⁰:

A finalidade funda-se no fato do conhecimento causal do homem, das possíveis consequências de sua atividade em determinada extensão, dispondo, assim, de diversos fins, e dirigindo o planejamento de sua atividade para obtenção desses fins. A pura causalidade não é dirigida a um fim, mas simples resultante fortuita de precedentes componentes causais. Figuradamente falando, a finalidade “vê”; a causalidade é “cega”. [...]

Por essa razão, a “ação” – que deve ser típica, ilícita e culpável – deve ser dirigida a um fim, de modo que verificada a distinção aos acontecimentos puramente causais, os quais se devem a forças não governadas pela vontade. Exige-se, nesse sentido, um comportamento exterior e a vontade dirigida a um fim, do ponto de vista subjetivo. A partir da teoria finalista, o dolo e culpa passaram a integrar a conduta, compondo o fato típico, e não mais a culpabilidade, como anteriormente²⁰¹.

Dessa forma, entendemos que a verificação da possibilidade – ou impossibilidade – de responsabilização penal dos sistemas de inteligência artificial pela divulgação de *fake news* e por crimes de ódio praticados por sistemas de inteligência artificial perpassa necessariamente pela análise dos elementos do crime adotados pela dogmática jurídico-penal brasileira, conforme trataremos em seguida.

4.3.2 A conduta

A *conduta*, também definida como ação ou ato²⁰², corresponde ao primeiro elemento da estrutura analítica do delito. Se inexistente conduta, não haverá crime, de igual sorte. Nesse sentido, convém apresentarmos algumas concepções doutrinárias acerca do que corresponde à noção de conduta.

De início, Ângelo Roberto Ilha da Silva, em que pese reconheça inexistir um

²⁰⁰ HUNGRIA, Néelson; FRAGOSO, Heleno Cláudio. **Comentários ao Código Penal, volume I, tomo II: arts. 11 ao 27**. 6. ed. Rio de Janeiro: Forense, 1983. p. 405.

²⁰¹ SILVA, Thiago Vítor Lima da; BEZERRA, Germano Cunha. **Responsabilidade penal nos delitos de inteligência artificial**. Revista Eletrônica Nova Roma: v. 2, n. 1, 2018. Disponível em: <https://faculdadenovaroma.com.br/revista/revistas/escola-de-direito-2.pdf#page=80>. Acesso em: 03 ago. 2023. p. 90.

²⁰² A propósito, Eugenio Raúl Zaffaroni e José Henrique Pierangeli entendem que prescindível a diferenciação entre “ato”, “ação” ou “conduta”, na medida em que “a consideração a respeito da extensão da atividade é determinada pelo tipo e não pela expressão que usamos no nível ôntico-ontológico.”. ZAFFARONI, Eugenio Raúl; PIERANGELI, José Henrique. **Manual de direito penal brasileiro: parte geral**. 2. ed. São Paulo: Revista dos Tribunais, 1999. p. 412.

sistema da teoria do delito perfeito, entende que o sistema mais coerente de todos é o finalista²⁰³, prevalente com a Reforma do Código Penal em 1984²⁰⁴, o qual estabelece “um sistema apto para tomarmos como ponto de partida”²⁰⁵. Segundo o finalismo de Hans Welzel²⁰⁶:

A ação é, portanto, um acontecimento final e não puramente causal. A finalidade, o caráter final da ação, baseia-se no fato de que o homem, graças ao seu saber causal, pode prever, dentro de certos limites, as possíveis consequências de sua conduta, designar-lhe fins diversos e dirigir sua atividade, conforme um plano, à consecução desses fins.

Nesse sentido, segundo Ilha da Silva, “tendo em conta o fato de a pessoa, dentro de certos limites, poder prever as consequências da ação humana, o resultado causal é governando pela vontade”²⁰⁷. Na mesma linha, para Cezar Roberto Bitencourt²⁰⁸, ao se adotar o finalismo na conceituação da ação humana relevante para o Direito Penal, somente pode vir a constituir um ilícito penal a manifestação exterior da ação humana voluntária, guiada por uma finalidade.

Guilherme de Souza Nucci, refere que, à luz do finalismo, “conduta é a ação ou omissão, voluntária e consciente, implicando um comando de movimentação ou inércia do corpo humano, voltado a uma finalidade”²⁰⁹. Por essa razão, “para a caracterização da conduta é indispensável a existência do binômio vontade e consciência”²¹⁰. A vontade corresponde ao querer ativo, que é apto a levar o ser humano a praticar um ato, livremente, enquanto a consciência diz respeito à possibilidade que o ser humano possui de separar o mundo que o cerca dos próprios atos, realizando um julgamento moral das suas atitudes²¹¹. Heleno Fragozo define

²⁰³ SILVA, Ângelo Roberto Ilha da. **Curso de direito penal: parte geral**. Porto Alegre: Livraria do Advogado. 2020. p. 208.

²⁰⁴ *Ibidem*, p. 203.

²⁰⁵ *Ibidem*, p. 209.

²⁰⁶ WELZEL, Hans. **O novo sistema jurídico-penal: uma introdução à doutrina da ação finalista**. Tradução de: Luiz Regis Prado. São Paulo: Revista dos Tribunais, 2015. *E-book*. Disponível em: <https://proview.thomsonreuters.com/launchapp/title/rt/monografias/103790584/v4/document/103909793>. Acesso em: 26 ago. 2023. n. p.

²⁰⁷ SILVA, Ângelo Roberto Ilha da. **Curso de direito penal: parte geral**. Porto Alegre: Livraria do Advogado. 2020. p. 208.

²⁰⁸ BITENCOURT, Cezar Roberto. **Tratado de direito penal: parte geral (arts. 1º a 120)**. v. 1. 29. ed. São Paulo: SaraivaJur, 2023. *E-book*. Disponível em: <https://app.minhabiblioteca.com.br/#!/books/9786553627109>. Acesso em: 03 ago. 2023. p. 148.

²⁰⁹ NUCCI, Guilherme de Souza. **Curso de direito penal: parte geral: arts. 1º a 120 do Código Penal**. 7. ed. Rio de Janeiro: Forense, 2023. *E-book*. Disponível em: <https://app.minhabiblioteca.com.br/#!/books/9786559646852>. Acesso em: 03 ago. 2023. p. 323.

²¹⁰ *Ibidem*, p. 324.

²¹¹ *Ibidem*.

ação como a “atividade humana conscientemente dirigida a um fim”²¹², o que se distingue “dos acontecimentos puramente causais, que se devem a forças não governadas pela vontade.”²¹³. Assim, a ação requer um comportamento exterior, e a vontade dirigida a um fim, do ponto de vista subjetivo.

Juarez Cirino dos Santos²¹⁴ sustenta que:

Considerando as funções teóricas, metodológicas e práticas do conceito de ação, definido *casualmente* como causação de resultado exterior por comportamento humano voluntário, *teleologicamente* como realização de atividade final, [...] é possível concluir que a definição capaz de identificar o traço mais específico e, ao mesmo tempo, a característica mais geral da ação humana, parece ser a definição do modelo *teleológico* de ação. A definição de ação como *atividade dirigida pelo fim* (nobre ou abjeto, altruísta ou egoísta, legal ou criminoso) destaca o traço que diferencia a ação de todos os demais fenômenos humanos e naturais, e permite delimitar a base real capaz de incorporar os atributos axiológicos do conceito de crime, como ação tipicamente injusta e culpável.

Por sua vez, para Juarez Tavares²¹⁵:

Conduta humana é, antes de tudo, uma conduta social, na qual devam estar presentes todos os elementos que congreguem a relação do sujeito com os demais, mediante a elaboração e consecução de um processo de comunicação. A pessoa humana, ao realizar uma conduta e orientá-la sob o âmbito da norma penal, deve ser dotada também das condições de fazê-lo em função de outrem e, portanto, de ser não apenas causadora de efeitos, mas de dirigir conscientemente o processo causal com vistas a esses efeitos. Daí se dizer que a conduta humana é uma conduta performática.

Importante registrar, contudo, que Eugenio Raúl Zaffaroni, Nilo Batista, Alejandro Alagia e Alejandro Slokar²¹⁶ sustentam a inexistência de um conceito ôntico de ação, razão pela qual existiriam, ao contrário, limites ônticos à construção jurídico-penal do conceito de ação. Ainda assim, a partir das definições carreadas, verificamos que há dois denominadores comuns identificáveis a todos os conceitos de conduta/ação.

Em primeiro lugar, é necessário que a conduta seja humana. Como bem destacam Eugenio Raúl Zaffaroni e José Henrique Pierangeli, revisando o Código

²¹² HUNGRIA, Nélon; FRAGOSO, Heleno Cláudio. **Comentários ao Código Penal, volume I, tomo II: arts. 11 ao 27**. 6. ed. Rio de Janeiro: Forense, 1983. p. 407.

²¹³ *Ibidem*.

²¹⁴ SANTOS, Juarez Cirino dos. **Direito penal: parte geral**. 8. ed. Florianópolis: Tirant lo Blanch, 2018. p. 124-125.

²¹⁵ TAVARES, Juarez. **Teoria do Injusto Penal**. 4. ed. São Paulo: Tirant lo Blanch, 2019. p. 195.

²¹⁶ ZAFFARONI, Eugenio Raúl; BATISTA, Nilo; ALAGIA, Alejandro; SLOKAR, Alejandro. **Direito penal brasileiro, segundo volume: teoria do delito: introdução histórica e metodológica, ação e tipicidade**. 2. ed. Rio de Janeiro: Revan, 2010. p. 101.

Penal, veremos que suas disposições se referem a condutas humanas²¹⁷. Da mesma forma, René Ariel Dotti assevera que o ilícito penal é fruto exclusivo da conduta humana²¹⁸. Em segundo lugar, é necessário que a conduta esteja dirigida e pautada em um fim.

Constata-se, então, que a teoria da ação adotada pelo nosso ordenamento jurídico é a finalista, de modo que a vontade consciente do seu fim é requisito para imputação de um delito segundo a teoria²¹⁹. Sem vontade não há, pois, conduta²²⁰. Deve-se, então, para a verificação da existência ou não de um delito, analisar se se trata de ação ou omissão voluntária direcionada a um fim, de modo que, não sendo, torna-se prescindível a verificação dos demais elementos da estrutura analítica do crime.

Veja-se, a propósito, que os dois elementos da conduta – necessidade de ação humana e vontade dirigida a uma finalidade – estão intimamente relacionados, na medida em que se considera que apenas o ser humano teria capacidade de agir com consciência do seu fim²²¹. Como vimos, ainda que um sistema de inteligência artificial possa proferir autônoma e independentemente discurso de ódio ou divulgar informações falsas, tais fatos estão associados à própria estrutura do sistema de IA, que “aprende” conforme a interação dos usuários e com as informações constantes de um banco de dados provenientes da internet.

Mediante esse raciocínio, o sistema Tay da Microsoft, por exemplo, cujo caso foi analisado anteriormente, não teria cometido um crime, na medida em que:

Ao publicar o comentário de cunho racista – por decisão própria – dirigido à pessoa do presidente, Tay infringiria, à luz do ordenamento jurídico brasileiro, uma norma penal. Contudo, não cometeu crime, já que não desempenhou uma ação, [...] porquanto não é um ser humano dotado de razão,

²¹⁷ ZAFFARONI, Eugenio Raúl; PIERANGELI, José Henrique. **Manual de direito penal brasileiro: parte geral**. 2. ed. São Paulo: Revista dos Tribunais, 1999. p. 409.

²¹⁸ DOTTI, Rene Ariel. **Curso de direito penal: parte geral**. Rio de Janeiro: Forense, 2001. p. 302.

²¹⁹ SILVA, Thiago Vitor Lima da; BEZERRA, Germano Cunha. **Responsabilidade penal nos delitos de inteligência artificial**. Revista Eletrônica Nova Roma: v. 2, n. 1, 2018. Disponível em: <https://faculdadenovaroma.com.br/revista/revistas/escola-de-direito-2.pdf#page=80>. Acesso em: 03 ago. 2023. p. 94.

²²⁰ ZAFFARONI, Eugenio Raúl; PIERANGELI, José Henrique. **Manual de direito penal brasileiro: parte geral**. 2. ed. São Paulo: Revista dos Tribunais, 1999. p. 413.

²²¹ SILVA, Thiago Vitor Lima da; BEZERRA, Germano Cunha. **Responsabilidade penal nos delitos de inteligência artificial**. Revista Eletrônica Nova Roma: v. 2, n. 1, 2018. Disponível em: <https://faculdadenovaroma.com.br/revista/revistas/escola-de-direito-2.pdf#page=80>. Acesso em: 03 ago. 2023. p. 85.

inadmissível, por esse motivo, pela teoria finalista²²².

Com efeito, a conjugação dos dois elementos leva à conclusão de que, à luz da estrutura analítica, tripartida e finalista do crime, não é possível a responsabilização penal dos sistemas de IA. Assim, a “conduta” *latu sensu* realizada – isto é, os efeitos produzidos no mundo real a partir das decisões tomadas pelos sistemas – não possui dimensão penal, tendo em vista que inexistente vontade consciente dirigida a uma finalidade.

4.3.3 A relação de causalidade

Ao analisarmos a possibilidade de o sistema de inteligência artificial responder penalmente, ele próprio, pelo ilícito penal, constatamos, examinando os conceitos de *conduta* e *ação*, a inviabilidade de tal solução. Nesse contexto, passa-se ao questionamento acerca da eventual possibilidade de a pessoa física – o usuário e, especialmente, o programador – responder pelo crime cometido autônoma e independentemente pelo sistema de IA.

Tal posição é, a propósito, sustentada pelo jurista espanhol Carlos Casabona²²³. Segundo ele, “ao não aplicar o conceito de culpa a um robô ou sistemas inteligentes, porque a culpa é inerente aos seres humanos e não a máquinas, abre-se uma nova discussão para a doutrina jurídica [...]”. Nesse cenário, entende que quem deveria responder por este tipo de crime seriam os seres humanos que participaram do desenvolvimento do software de IA.

Com a devida vênia, discordamos da posição do doutrinador, que não se sustentaria na dogmática jurídico-penal brasileira. Com efeito, entendemos que a resolução do questionamento perpassa, necessariamente, pela análise da *relação de causalidade* e, conseqüentemente, pelo disposto no artigo 13 do Código Penal.

A *relação de causalidade*, segundo define Luiz Regis Prado²²⁴, corresponde a

²²² *Ibidem*, p. 101.

²²³ PUCRS. **Inteligência artificial pode ter responsabilidade penal**. 2018. Disponível em: <https://www.pucrs.br/blog/inteligencia-artificial-pode-ter-responsabilidade-penal>. Acesso em: 03 ago. 2023.

²²⁴ PRADO, Luiz Regis. **Curso de direito penal brasileiro: volume 1**. 6. ed. São Paulo: Revista dos Tribunais, 2018. *E-book*. Disponível em: <https://proview.thomsonreuters.com/launchapp/title/rt/monografias/122952706>. Acesso em: 04 ago. 2023. n. p.

“[...] um liame causal natural – prévio –, e não pode ser substituído por nenhum critério normativo. Noutra dizer: apresenta-se como vínculo material entre a ação humana como causa, e determinado evento como seu efeito.”. Segundo prevê o artigo 13 do Código Penal: “O resultado, de que depende a existência do crime, somente é imputável a quem lhe deu causa. Considera-se causa a ação ou omissão sem a qual o resultado não teria ocorrido.”.

Nesse sentido, da redação do dispositivo legal é possível extrair duas conclusões²²⁵: em primeiro lugar, sem causalidade não há imputação, tendo em vista o princípio da responsabilidade pessoal, em oposição à responsabilidade solidária, existente no direito privado; em segundo lugar, em se admitindo a causalidade na omissão, o processo de imputação também se estende ao risco de produção do resultado em face da ação devida.

Com a redação do artigo 13, verificamos que o Código Penal adotou a chamada *teoria da equivalência dos antecedentes*, ou teoria da *conditio sine qua non*, a qual considera que “qualquer das condições que compõem a totalidade dos antecedentes é causa do resultado, pois a sua inocorrência impediria a produção do evento”²²⁶. Nesse sentido, utiliza-se da fórmula do “processo ou método indutivo hipotético de eliminação”, na qual considera-se causa de um resultado toda condição que, se suprimida mentalmente, levaria ao desaparecimento do resultado²²⁷.

Essa interpretação permite chegar à conclusão de que, em um caso de homicídio, a venda de arma ou mesmo a fabricação de uma arma, ainda que em atividade lícita, seriam consideradas causas do resultado. Isso porque, sem a arma, não teria havido os tiros, os quais, em seguida, levaram à morte.

A teoria, contudo, como assevera Nélson Hungria²²⁸, não pode ser o único pressuposto da punibilidade, porquanto é subordinada à culpabilidade do agente, de

²²⁵ TAVARES, Juarez. **Teoria do Injusto Penal**. 4. ed. São Paulo: Tirant lo Blanch, 2019. p. 263.

²²⁶ NUCCI, Guilherme de Souza. **Curso de direito penal: parte geral: arts. 1º a 120 do Código Penal**. 7. ed. Rio de Janeiro: Forense, 2023. *E-book*. Disponível em: <https://app.minhabiblioteca.com.br/#/books/9786559646852>. Acesso em: 04 ago. 2023. p. 363.

²²⁷ PRADO, Luiz Regis. **Curso de direito penal brasileiro: volume 1**. 6. ed. São Paulo: Revista dos Tribunais, 2018. *E-book*. Disponível em: <https://proview.thomsonreuters.com/launchapp/title/rt/monografias/122952706>. Acesso em: 04 ago. 2023. n. p.

²²⁸ HUNGRIA, Nélson; FRAGOSO, Heleno Cláudio. **Comentários ao Código Penal, volume I, tomo II: arts. 11 ao 27**. 6. ed. Rio de Janeiro: Forense, 1983. p. 51.

modo que, após a averiguação de um evento penalmente típico em sua objetividade, é necessário verificar não somente se foi causado por alguém, mas, também, se o agente procedeu dolosa ou culposamente. Tal requisito corresponde, pois, a um corretivo à amplitude excessiva do conceito de causa.

O Código Penal, a propósito, no § 1º do artigo 13, dispõe que: “A superveniência de causa relativamente independente exclui a imputação quando, por si só, produziu o resultado; os fatos anteriores, entretanto, imputam-se a quem os praticou.”. O artigo, ao dispor a respeito da interrupção da causalidade, buscou limitar o regresso infinito do processo causal²²⁹. Tais causas relativamente independentes surgem ligadas às causas geradas pelo agente, mas possuem força suficiente para gerar o resultado por si mesmas²³⁰, de modo que o agente responderá somente pelo que já praticou.

Ademais, há a teoria da causalidade adequada, a qual serve de critério delimitador da causalidade, a qual considera que a condição é causa quando se apresenta geralmente proporcionada ou adequada ao resultado, constatado por meio de um juízo de probabilidade. Assim, verificar-se-á se há um nexos causal ligando a ação do agente como causa ao resultado como efeito²³¹. Um evento somente será considerado causa do evento delituoso quando for apto e idôneo a gerar o resultado.

A partir desta análise, entendemos que, ao se adotar como válida a premissa de que a pessoa física – em especial a figura do programador – deve ser considerada penalmente responsável pelos atos praticados por sistemas de inteligência artificial, se estaria permitindo indevido regresso para além da relação direta de causalidade. Em verdade, inexistente relação de causalidade entre a atividade do programador que desenvolveu o sistema de IA e o ilícito penal eventualmente praticado.

Como vimos, ainda que o Código Penal adote a chamada *teoria da equivalência dos antecedentes*, é necessária a imposição de restrições para, justamente, evitar regresso *ad infinitum*. Por essa razão, compreendemos que a responsabilização da pessoa física por atos independentes praticados pelos sistemas de IA guarda correspondência à hipótese da indevida responsabilização do fabricante

²²⁹ TAVARES, Juarez. **Teoria do Injusto Penal**. 4. ed. São Paulo: Tirant lo Blanch, 2019. p. 270.

²³⁰ NUCCI, Guilherme de Souza. **Curso de direito penal: parte geral: arts. 1º a 120 do Código Penal**. 7. ed. Rio de Janeiro: Forense, 2023. *E-book*. Disponível em: <https://app.minhabiblioteca.com.br/#/books/9786559646852>. Acesso em: 04 ago. 2023. p. 373.

²³¹ DOTTI, Rene Ariel. **Curso de direito penal: parte geral**. Rio de Janeiro: Forense, 2001. p. 319.

ou do vendedor de uma arma pelo homicídio praticado por terceiro.

Outrossim, Lucas Hinckel Teider e Gabriel Pivatto dos Santos²³² defendem a possibilidade de responsabilização da pessoa física mediante a aplicação da teoria da imputação objetiva, senão vejamos:

Na eventualidade de desejo político-criminal de imputação de fato criminalmente punível ao indivíduo nos casos de conduta praticada por inteligência artificial, possibilidade a ser cogitada seria a adoção e aplicação da teoria da imputação objetiva, devendo o tipo objetivo conglobar a causalidade material e a normativa.

Com efeito, a teoria da imputação objetiva é originária dos trabalhos do civilista Karl Larenz, trasladada à seara penal por Richard Honig, que, posteriormente, na década de 1970, ganhou impulso com Claus Roxin²³³. Pressupõe a imputação objetiva que, para que alguém seja penalmente responsabilizado penalmente por uma conduta que desenvolveu, é necessária “a criação ou o incremento de um perigo juridicamente intolerável e não permitido ao bem jurídico protegido, bem como a concretização desse perigo em resultado típico”²³⁴.

Trata-se a imputação objetiva de um critério normativo-restritivo da *causalidade natural*, na medida em que considera que nem todo resultado causado é imputável à ação do agente. Assim, a causação de um resultado típico somente realizará o tipo objetivo delito na hipótese de o agente ter criado um perigo juridicamente desaprovado que levou ao resultado²³⁵. Por essa razão, se a ação do autor não cria risco do resultado, ou se o risco criado pela ação do autor não se realiza no resultado, então o resultado não pode ser imputado ao autor²³⁶.

Verificamos, então, que a imputação objetiva corresponde a um segundo

²³² TEIDER, Lucas Hinckel; SANTOS, Gabriel Pivatto dos. **Responsabilidade penal da inteligência artificial (?): a problemática relacionada ao elemento da conduta na clássica estrutura analítica do delito**. In: (Coord.) GONZÁLEZ, Javier Garcia; LOZANO, Álvaro Alzina; RODRÍGUEZ, Gabriel Martín; GONÇALVES, Rubén Miranda; VEIGA, Fábio da Silva. *El derecho público y privado ante las nuevas tecnologías*. Madrid: Dykinson, 2020. Disponível em: <https://doi.org/10.2307/j.ctv153k4rn>. Acesso em: 04 jul. 2023. p. 192.

²³³ DOTTI, Rene Ariel. **Curso de direito penal: parte geral**. Rio de Janeiro: Forense, 2001. p. 364.

²³⁴ *Ibidem*.

²³⁵ PRADO, Luiz Regis. **Curso de direito penal brasileiro: volume 1**. 6. ed. São Paulo: Revista dos Tribunais, 2018. *E-book*. Disponível em: <https://proview.thomsonreuters.com/launchapp/title/rt/monografias/122952706>. Acesso em: 04 ago. 2023. n. p.

²³⁶ SANTOS, Juarez Cirino dos. **Direito penal: parte geral**. 8. ed. Florianópolis: Tirant lo Blanch, 2018. p. 137.

patamar de análise da responsabilização penal. Nesse sentido, eventualmente, pode-se ver reconhecida a relação de causalidade, mas, posteriormente, com base nos critérios da imputação objetiva, irá se constatar a impossibilidade de responsabilização.

Por essa razão, na linha do que sustentam Teider e Santos²³⁷, para responsabilização da pessoa física pela conduta praticada pelo ato realizado pelo sistema de inteligência artificial, é necessário que seja criado um risco não permitido para aplicação da teoria da imputação objetiva. Além disso, impositiva a percepção das circunstâncias de realização do risco no resultado concreto, bem como o fato de que o resultado deve se encontrar dentro do alcance do próprio tipo.

Portanto, entendemos que, via de regra, não se faz possível a responsabilização da pessoa física – especialmente o programador – por ilícitos penais praticados pelo sistema de IA com base na teoria da imputação objetiva. Isso porque “não vale para a imputação qualquer risco, mas apenas o risco indevido e materializado na conduta e no resultado típicos”²³⁸. Sendo o risco permitido, não haverá imputação. Avaliamos que, para aplicação da teoria, se faz necessária a constatação de um risco que transcenda ao ordinário e meramente aos riscos que o desenvolvimento de um sistema tecnológico representa, sob pena de responsabilização objetiva.

4.3.4 A responsabilização penal da empresa desenvolvedora

Doravante, ao se falar na responsabilização penal pela prática de crimes por sistemas de IA, passa-se a cogitar na eventual possibilidade de responsabilizar penalmente a pessoa jurídica, mais especificamente, a empresa desenvolvedora do sistema. Em se admitindo como correta e apropriada tal tese, poderíamos falar, portanto, a título de exemplo, que a empresa OpenAI, desenvolvedora do ChatGPT, na hipótese de o *chat* proferir discurso de ódio, ou ainda a Microsoft, no caso do robô Tay, de 2016, se ocorrido no Brasil, seriam penalmente responsáveis pelos atos praticados pelos sistemas. Nesse sentido, em seguida, passaremos à análise da

²³⁷ TEIDER, Lucas Hinckel; SANTOS, Gabriel Pivatto dos. **Responsabilidade penal da inteligência artificial (?): a problemática relacionada ao elemento da conduta na clássica estrutura analítica do delito**. In: (Coord.) GONZÁLEZ, Javier Garcia; LOZANO, Álvaro Alzina; RODRÍGUEZ, Gabriel Martín; GONÇALVES, Rubén Miranda; VEIGA, Fábio da Silva. *El derecho público y privado ante las nuevas tecnologías*. Madrid: Dykinson, 2020. Disponível em: <https://doi.org/10.2307/j.ctv153k4rn>. Acesso em: 04 jul. 2023. p. 192.

²³⁸ TAVARES, Juarez. **Teoria do Injusto Penal**. 4. ed. São Paulo: Tirant lo Blanch, 2019. p. 297.

problemática.

Conforme destaca Alamiro Velludo Salvador Netto²³⁹:

[...] a responsabilização penal da pessoa jurídica talvez seja o maior desafio imposto à dogmática penal nos últimos anos. A mudança dos destinatários das normas jurídicas exige a revisitação, e conseqüente representar, de praticamente todos os institutos da teoria do delito e da teoria da pena, sempre com vistas a não romper, por um lado, com a identidade nuclear e garantista do sistema criminal e, por outro, proporcional o mínimo de eficiência e capacidade de tutela sobre as atividades empresariais.

A responsabilização da pessoa jurídica no direito penal brasileiro, mesmo positivada na Lei 9.605/98, é tema de grande controvérsia na doutrina: de um lado, há aqueles que defendem a responsabilização penal da pessoa jurídica, como Alamiro Velludo Salvador Netto²⁴⁰ e Rodrigo Andrade Viviani²⁴¹, de outro, há aqueles que são contrários, como Eugenio Raúl Zaffaroni e José Henrique Pierangeli²⁴², Juarez Cirino dos Santos²⁴³ e Heloisa Estellita²⁴⁴. Sem pretender proceder à exaustiva análise dos argumentos favoráveis e contrários, realizaremos breve digressão acerca do panorama da matéria na doutrina brasileira.

Com efeito, Eugenio Raúl Zaffaroni e José Henrique Pierangeli²⁴⁵ são categóricos ao afirmar que “a pessoa jurídica não pode ser autora de delito, porque não tem capacidade de conduta humana no seu sentido ôntico-ontológico”. Por essa razão, sustentam que a ausência de capacidade de conduta na pessoa consiste em fator central ao analisar a problemática, destacando que negar considerar a pessoa jurídica como autora do delito não implica na impossibilidade de os diretores e administradores serem punidos, bem como na impossibilidade de aplicação de

²³⁹ NETTO, Alamiro Velludo Salvador. **Responsabilidade penal da pessoa jurídica**. 3. ed. São Paulo: Thomson Reuters Brasil, 2023. *E-book*. Disponível em: <https://proview.thomsonreuters.com/launchapp/title/rt/monografias/148836979>. Acesso em: 06 ago. 2023.

²⁴⁰ *Ibidem*.

²⁴¹ VIVIANI, Rodrigo Andrade. **Responsabilidade penal da pessoa jurídica: aspectos controvertidos no direito brasileiro**. Curitiba: Juruá, 2008.

²⁴² ZAFFARONI, Eugenio Raúl; PIERANGELI, José Henrique. **Manual de direito penal brasileiro: parte geral**. 2. ed. São Paulo: Revista dos Tribunais, 1999. p. 409-411.

²⁴³ SANTOS, Juarez Cirino dos. **Direito penal: parte geral**. 8. ed. Florianópolis: Tirant lo Blanch, 2018. p. 717-749.

²⁴⁴ ESTELLITA, Heloisa. **A responsabilidade isolada da pessoa jurídica por crimes ambientais**. 2014. Disponível em: <https://www.conjur.com.br/2014-dez-16/heloisa-estellita-responsabilidade-isolada-crimes-ambientais#author>. Acesso em: 06 ago. 2023.

²⁴⁵ ZAFFARONI, Eugenio Raúl; PIERANGELI, José Henrique. **Manual de direito penal brasileiro: parte geral**. 2. ed. São Paulo: Revista dos Tribunais, 1999. p. 410.

sanções administrativas à pessoa jurídica.

Juarez Cirino dos Santos, de igual sorte, reputa “absurda” a tese da responsabilidade penal da pessoa jurídica²⁴⁶. Nesse sentido, os principais argumentos expostos pelo doutrinador podem ser assim sintetizados²⁴⁷:

I. Incapacidade de ação, na medida em que, como fundamento psicossomático do crime, designa fenômeno exclusivamente humano.

II. Incapacidade de dolo, isto é, a “vontade coletiva” da pessoa jurídica é incapaz de realizar um crime.

III. Incapacidade de imprudência, posto que consistente em fenômeno fundado na capacidade individual, impossibilitando a analogia do “homem prudente” para o conceito de “empresa prudente”.

IV. Incapacidade de omissão, considerando que, de igual sorte, é incapaz de ação.

V. Incapacidade de culpabilidade, porque as condições de maturidade e de sanidade mental dos seres humanos não existem na entidade incorpórea da pessoa jurídica, bem como é impossível de considerar a possibilidade de agir de outro modo e a possibilidade de a pessoa jurídica representar a natureza proibida do tipo de injusto.

VI. Incompatibilidade dos objetivos da pena criminal com a pessoa jurídica, representados na retribuição da culpabilidade e na prevenção da criminalidade.

Heloisa Estellita²⁴⁸, na mesma esteira, assevera que a responsabilização da pessoa jurídica é injustificável sob o fundamento da “impunidade”, sustentando a possibilidade de se estar abrindo mão dos pressupostos de um “direito penal de culpabilidade” e de estar se desenvolvendo um “sistema penal paralelo”:

Admitir pura e simplesmente a acusação contra uma pessoa jurídica abrindo-se mão dos pressupostos de um direito penal de culpabilidade (em sentido

²⁴⁶ SANTOS, Juarez Cirino dos. **Direito penal: parte geral**. 8. ed. Florianópolis: Tirant lo Blanch, 2018. p. 723.

²⁴⁷ SANTOS, Juarez Cirino dos. **Direito penal: parte geral**. 8. ed. Florianópolis: Tirant lo Blanch, 2018. p. 728-749.

²⁴⁸ ESTELLITA, Heloisa. **A responsabilidade isolada da pessoa jurídica por crimes ambientais**. 2014. Disponível em: <https://www.conjur.com.br/2014-dez-16/heloisa-estellita-responsabilidade-isolada-crimes-ambientais#author>. Acesso em: 06 ago. 2023.

amplo) é, ao mesmo tempo, negar vigência a todos os dispositivos penais, constitucionais ou infralegais, que atrelam a responsabilidade penal a uma conduta pessoal e subjetiva. Se é possível responsabilizar a pessoa jurídica sem comprovada culpabilidade (ainda que adaptada a tal “ator”), o que impedirá de se fazer o mesmo quanto à pessoa física?

É essa a lógica nefasta que perturba na admissão pura e simples do processamento contra a pessoa jurídica sem que estabeleçam, antes, os pressupostos de sua culpabilidade. Ou seja, sem que se adaptem — se é que isto é mesmo possível — a essa “pessoa” os pressupostos elementares da responsabilidade penal, assentados na prática de uma conduta típica, ilícita e culpável, no sentido de reprovação pessoal por não agir conforme o direito, quando, nas circunstâncias lhe era possível assim proceder. [...]

A despeito das críticas, considera-se que, no cenário brasileiro, a responsabilidade penal da pessoa jurídica surgiu a partir da Constituição Federal de 1988, por meio de dois dispositivos, quais sejam, o artigo 173, § 5º – alocado no Título VII (“Da ordem econômica”), Capítulo I (“Dos princípios gerais da atividade econômica”) – e o artigo 225, § 3º – constante no Título VIII (“Da ordem social”), Capítulo VI (“Do meio ambiente”). Vejamos a redação dos dispositivos:

Art. 173, § 5º A lei, sem prejuízo da responsabilidade individual dos dirigentes da pessoa jurídica, estabelecerá a responsabilidade desta, sujeitando-a às punições compatíveis com sua natureza, nos atos praticados contra a ordem econômica e financeira e contra a economia popular.

Art. 225, § 3º As condutas e atividades consideradas lesivas ao meio ambiente sujeitarão os infratores, pessoas físicas ou jurídicas, a sanções penais e administrativas, independentemente da obrigação de reparar os danos causados.

Ao analisarmos os dispositivos, verificamos que o artigo 173, § 5º, que trata sobre “atos praticados contra a ordem econômica e financeira e contra a economia popular”, não estabeleceu expressamente a responsabilidade penal da pessoa jurídica. Em verdade, trata da “responsabilidade” *lato sensu*, e não em responsabilidade criminal, deixando margem para dúvidas²⁴⁹.

No entanto, setor da doutrina entende que a intenção do constituinte era a atribuição de responsabilização penal às condutas descritas no artigo. Rodrigo Andrade Viviani²⁵⁰ entende que, malgrado não regulado infraconstitucionalmente, “nada obsta que uma lei infraconstitucional regule o artigo 173, § 5º, firmando a citada responsabilidade, pois o intuito do constituinte foi a de penalizar a pessoa

²⁴⁹ VIVIANI, Rodrigo Andrade. **Responsabilidade penal da pessoa jurídica: aspectos controversos no direito brasileiro**. Curitiba: Juruá, 2008. p. 80.

²⁵⁰ *Ibidem*.

jurídica, ainda que ele não tenha mencionado explicitamente em qual esfera.”.

Por sua vez, o artigo 225, § 3º restou regulamentado pela Lei n.º 9.605/98, conhecida como Lei de Crimes Ambientais. Segundo o artigo 3º, caput, da Lei:

Art. 3º As pessoas jurídicas serão responsabilizadas administrativa, civil e penalmente conforme o disposto nesta Lei, nos casos em que a infração seja cometida por decisão de seu representante legal ou contratual, ou de seu órgão colegiado, no interesse ou benefício da sua entidade.

A respeito do dispositivo, Viviani²⁵¹ comenta que existem dois requisitos para que a pessoa jurídica possa ser responsabilizada criminalmente pela prática de crimes ambientais: “1) que a infração seja cometida por decisão de seu representante legal ou contratual, ou de seu órgão colegiado;” e “2) que o crime praticado seja em benefício ou interesse da entidade coletiva.”. De igual sorte, avalia que, detendo a pessoa jurídica capacidade de ação em outros ramos jurídicos, como nas esferas cível e administrativa, nada impede que o direito penal reformule seus conceitos para atribuir à pessoa jurídica ação própria para o cometimento de crimes, com os consectários do dolo, da culpa e dos elementos de culpabilidade²⁵².

Acerca da eventual necessidade de persecução penal da pessoa física em conjunto, insta salientar, ainda, que o Supremo Tribunal Federal, no julgamento do paradigmático Recurso Extraordinário n.º 548.181/PR, de Relatoria da Ministra Rosa Weber²⁵³, decidiu que:

[...] O art. 225, § 3º, da Constituição Federal não condiciona a responsabilização penal da pessoa jurídica por crimes ambientais à simultânea persecução penal da pessoa física em tese responsável no âmbito da empresa. A norma constitucional não impõe a necessária dupla imputação.

Veja-se que, mesmo aqueles que defendem a possibilidade de responsabilização penal da pessoa jurídica, sustentam a necessidade de identificação de uma conduta praticada por um indivíduo. Nesse sentido, Alamiro Velludo Salvador Netto²⁵⁴ defende que:

²⁵¹ VIVIANI, Rodrigo Andrade. **Responsabilidade penal da pessoa jurídica: aspectos controvertidos no direito brasileiro**. Curitiba: Juruá, 2008. p. 100.

²⁵² *Ibidem*, p. 114.

²⁵³ BRASIL. Supremo Tribunal Federal. **Recurso Extraordinário n.º 548181**. Recorrente: Ministério Público Federal. Recorrido: Petróleo Brasileiro S/A – Petrobrás. Relatora: Ministra Rosa Weber. Primeira Turma. Julgado em 06/08/2013. Disponível em: <https://redir.stf.jus.br/paginadorpub/paginador.jsp?docTP=TP&docID=7087018>. Acesso em: 06 ago. 2023.

²⁵⁴ NETTO, Alamiro Velludo Salvador. **Responsabilidade penal da pessoa jurídica**. 3. ed. São Paulo:

[...] o sistema de heteroresponsabilidade pode até prescindir da identificação da pessoa física atuante, porém jamais superará a necessidade de imputação de uma conduta, específica e delineada, praticada por um indivíduo. De acordo com o próprio acórdão do STF, a conduta perpetrada pela pessoa natural a ser imputada à pessoa jurídica necessita ser descrita na denúncia, apontando-se expressamente todos os seus contornos objetivos e subjetivos. Isto é, qual o órgão que a realizou; se de forma dolosa ou culposa; se dentro das atribuições empresariais do realizador; e se em favor ou proveito do ente moral. Não ter eventualmente a identidade da pessoa física autora é algo diferente de não identificar a própria conduta realizada.

A partir dessa exposição, ao aproximar a análise à responsabilização penal da empresa desenvolvedora do sistema de inteligência artificial que comete um ilícito penal, podemos concluir, em síntese, pela sua impossibilidade. Alinhamo-nos com os fundamentos contrários à responsabilização penal, na medida em que:

[...] o Direito Penal não alcança nem o fornecedor/detentor nem o sistema autônomo, já que apenas o ser humano pode ser considerado sujeito ativo do crime e este [...] não desencadeia comportamento voluntário psiquicamente dirigido a um fim, muito menos age com culpa. Responsabilizar penalmente a empresa seria impossível considerando a teoria do crime e o atual entendimento da Suprema Corte [...] ²⁵⁵

Outrossim, é possível verificar que, até o momento, há apenas duas hipóteses de responsabilização penal da pessoa jurídica existentes do ordenamento jurídico brasileiro. A primeira, prevista no artigo 173, § 5º, da CF/88, ainda não regulamentado, e a segunda, constante no artigo 225, § 3º, da CF/88, regulamentado infraconstitucionalmente pela Lei n.º 9.605/98.

Nesse sentido, em que pese possa se cogitar em eventual responsabilização penal de pessoas jurídicas por outros injustos penais além daqueles previstos na CF/88, como entende Salvador Netto²⁵⁶ – ao contrário de Damásio Evangelista de Jesus²⁵⁷, que sustenta que a responsabilização estaria restrita às hipóteses previstas

Thomson Reuters Brasil, 2023. *E-book*. Disponível em: <https://proview.thomsonreuters.com/launchapp/title/rt/monografias/148836979>. Acesso em: 06 ago. 2023. p. RB-5.3.

²⁵⁵ SILVA, Thiago Vítor Lima da; BEZERRA, Germano Cunha. **Responsabilidade penal nos delitos de inteligência artificial**. Revista Eletrônica Nova Roma: v. 2, n. 1, 2018. Disponível em: <https://faculdadenovaroma.com.br/revista/revistas/escola-de-direito-2.pdf#page=80>. Acesso em: 03 ago. 2023. p. 100.

²⁵⁶ NETTO, Alamiro Velludo Salvador. **Responsabilidade penal da pessoa jurídica**. 3. ed. São Paulo: Thomson Reuters Brasil, 2023. *E-book*. Disponível em: <https://proview.thomsonreuters.com/launchapp/title/rt/monografias/148836979>. Acesso em: 06 ago. 2023. p. RB-5.1.

²⁵⁷ “Inaugurada, pois, nova fase no sistema jurídico de admissão da capacidade penal das pessoas coletivas, ainda que efetuada com a reserva de somente se mostrar cabível nas hipóteses indicadas na Constituição”. JESUS, Damásio de. Proteção ambiental e responsabilidade penal da pessoa jurídica. Carta Forense, São Paulo, n. 154, 2016. p. A7 *apud* NETTO, Alamiro Velludo Salvador. **Responsabilidade penal da pessoa jurídica**. 3. ed. São Paulo: Thomson Reuters Brasil, 2023. *E-*

na Constituição Federal –, ainda assim é necessário ter em mente que a responsabilização das pessoas jurídicas está associada à “delinquência empresarial”, isto é:

O posicionamento desses dispositivos nos aludidos títulos e capítulos demonstram coerência do legislador constitucional com a própria dimensão político-criminal da responsabilidade penal da pessoa jurídica. [...] o problema da responsabilização penal desses entes coletivos aparece fundamentalmente como um tema de Direito Penal Econômico, tendo em vista a capacidade que as empresas gradativamente assumiram de realizar condutas lesivas ao sistema tributário e financeiro, à concorrência, ao mercado de capitais, ao meio ambiente etc. Cuida-se, assim, de um novo ator apto a protagonizar agressões no âmbito do sistema criminal, principalmente em razão do papel central que ocupam na produção e circulação de riquezas, bem como dos efeitos dessas mesmas atividades na contemporaneidade²⁵⁸.

Tal conclusão, a nosso ver, decorre da interpretação hermenêutica dos dispositivos constitucionais supracitados, que, inequivocamente, associam a responsabilidade penal às peculiaridades da atividade empresarial. Entendemos que, ao se cogitar eventual responsabilização das pessoas jurídicas por crimes comuns, não associados à atividade empresarial, estar-se-ia meramente trasladando e estendendo a responsabilidade penal que seria atribuída à pessoa física para a pessoa jurídica, expandindo, desarrazoadamente, a dimensão do tipo penal.

Portanto, avaliamos que, atualmente, o direito penal brasileiro não permite a responsabilização penal da empresa desenvolvedora do sistema de inteligência artificial que comete um ilícito penal. Com efeito, uma dogmática jurídico-penal garantista e consonante à teoria do delito, concebido como uma conduta típica, ilícita e culpável, não deve permitir, a pretexto de suposta impunidade, a extensão da responsabilidade a quem, ao contrário, não pode ser responsabilizado.

4.3.5 Alternativas à responsabilização penal

Ao analisarmos a possibilidade de os sistemas de inteligência artificial e pessoas físicas e jurídicas a eles vinculadas responderem penalmente por ilícitos penais cometidos, chegamos à conclusão em sentido negativo, conforme expusemos

book. Disponível em: <https://proview.thomsonreuters.com/launchapp/title/rt/monografias/148836979>. Acesso em: 06 ago. 2023. p. RB-5.1.

²⁵⁸ NETTO, Alamiro Velludo Salvador. **Responsabilidade penal da pessoa jurídica**. 3. ed. São Paulo: Thomson Reuters Brasil, 2023. *E-book*. Disponível em: <https://proview.thomsonreuters.com/launchapp/title/rt/monografias/148836979>. Acesso em: 06 ago. 2023. p. RB-5.1.

alhures. Todavia, ainda que incabível a incidência do direito penal, entendemos que há outros ramos do direito que são igualmente capazes de atribuir resposta e sanção adequada a atos ilícitos cometidos pelos sistemas de IA.

Em que pese a impossibilidade de responsabilização penal, avaliamos cabível a incidência da responsabilidade civil. Consoante define Sérgio Cavalieri Filho²⁵⁹, “responsabilidade civil é um dever jurídico sucessivo que surge para recompor o dano decorrente da violação de um dever jurídico originário”.

Nesse sentido, conforme sustenta Cavalieri²⁶⁰, a ilicitude não é uma peculiaridade exclusiva ao direito penal, na medida em que, sendo ela compreendida como contrariedade entre a conduta e a norma jurídica, pode ter lugar em qualquer ramo do direito. José de Aguiar Dias²⁶¹, na mesma linha, destaca a possibilidade de repercussão de uma conduta no âmbito cível e/ou penal, alternativa ou concomitantemente:

Assim, certos fatos põem em ação somente o mecanismo recuperatório da responsabilidade civil; outros movimentam tão somente o sistema repressivo ou preventivo da responsabilidade penal; outros, enfim, acarretam, a um tempo, a responsabilidade civil e a penal, pelo fato de apresentarem, em relação a ambos os campos, incidência equivalente, conforme os diferentes critérios sob que entram em função os órgãos encarregados de fazer valer a norma respectiva.

Segundo o autor, o fundamento da responsabilidade civil e da responsabilidade penal são quase os mesmos, sendo diferentes as condições em que surgem, porque uma é mais exigente do que a outra quanto aos requisitos que devem coincidir para se efetivar²⁶². Nesse sentido, o ilícito penal difere-se do ilícito civil, o qual corresponde a um “*minus*” ou “*residum*”, de modo que as condutas humanas mais graves, capazes de atingir bens sociais de maior relevância, são sancionadas pela lei penal, enquanto a lei civil dá conta de reprimir condutas menos graves²⁶³.

²⁵⁹ CAVALIERI FILHO, Sérgio. **Programa de responsabilidade civil**. 16. ed. Barueri: Atlas, 2023. *E-book*. Disponível em: <https://app.minhabiblioteca.com.br/#/books/9786559775217>. Acesso em: 07 ago. 2023. p. 11.

²⁶⁰ *Ibidem*, p. 25.

²⁶¹ DIAS, José de Aguiar. **Da responsabilidade civil**. Rio de Janeiro: Forense, 1944. p. 11.

²⁶² *Ibidem*.

²⁶³ CAVALIERI FILHO, Sérgio. **Programa de responsabilidade civil**. 16. ed. Barueri: Atlas, 2023. *E-book*. Disponível em: <https://app.minhabiblioteca.com.br/#/books/9786559775217>. Acesso em: 07 ago. 2023. p. 25.

Por sua vez, Carlos Roberto Gonçalves²⁶⁴ destaca que, no caso da responsabilidade penal, o agente infringe uma norma de direito público, sendo lesado o interesse da sociedade, enquanto, na responsabilidade civil, o interesse diretamente lesado é o privado. De igual sorte, aponta que a responsabilidade penal é pessoal e intransferível, devendo o réu estar cercado de todas as garantias contra o Estado, que deve sempre arcar com o ônus probatório.

Ademais, no que diz respeito à tipicidade penal, é necessária a perfeita adequação do fato concreto ao tipo penal, enquanto, no âmbito cível, qualquer ação ou omissão pode gerar a responsabilidade civil, desde que viole direito e cause prejuízo a outrem²⁶⁵. A culpabilidade na área cível é, igualmente, mais ampla, ao passo que, na esfera criminal, é necessário certo grau ou intensidade. Ainda que a culpa civil e a culpa penal possuam os mesmos elementos, considera-se que o juízo criminal é mais exigente para aplicação da lei²⁶⁶.

Outra principal diferença da responsabilidade civil para a responsabilidade penal reside na possibilidade de, naquela, ser atribuída responsabilidade objetiva, isto é, a responsabilidade independentemente de culpa, fundamentada no risco. Com efeito, “na responsabilidade objetiva [...] é irrelevante o nexó psicológico entre o fato ou atividade e a vontade de quem a pratica, bem como o juízo de censura moral ou de aprovação da conduta.”²⁶⁷.

Segundo Judith Martins-Costa²⁶⁸, são pressupostos do dever de indenizar a antijuridicidade/a ilicitude da conduta do agente, o dano e o nexó causal a existir entre o dano e a conduta ilícita, isto é:

Antijuridicidade, dano e nexó de causalidade são, portanto, os três pressupostos gerais elementares, presentes tanto na responsabilidade contratual quanto na extracontratual; na responsabilidade subjetiva

²⁶⁴ GONÇALVES, Carlos Roberto. **Direito civil brasileiro, volume 4: responsabilidade civil**. 14. ed. São Paulo: Saraiva Educação, 2019. p. 43.

²⁶⁵ *Ibidem*, p. 44.

²⁶⁶ *Ibidem*.

²⁶⁷ CAVALIERI FILHO, Sérgio. **Programa de responsabilidade civil**. 16. ed. Barueri: Atlas, 2023. *E-book*. Disponível em: <https://app.minhabiblioteca.com.br/#/books/9786559775217>. Acesso em: 07 ago. 2023. p. 209.

²⁶⁸ MARTINS-COSTA, Judith. A linguagem da responsabilidade civil. In: (Coord.) BIANCHI, José Flávio; MENDONÇA PINHEIRO, Rodrigo Gomes de; ARRUDA ALVIM, Teresa. **Jurisdição e Direito Privado: Estudos em homenagem aos 20 anos da Ministra Nancy Andrighi no STJ**. São Paulo: Revista dos Tribunais, 2020. *E-book*. Disponível em: <https://proview.thomsonreuters.com/launchapp/title/rt/monografias/230263461>. Acesso em: 07 ago. 2023. p. RB-20.2

(informada pela culpa como fator de imputação) e na responsabilidade objetiva (informada pelo risco como fator de imputação)²⁶⁹.

Assim, a especificação do regime jurídico a que está sujeito o dever de indenizar – se o regime da responsabilidade subjetiva ou o da objetiva – obedece a determinados fatores de imputação: a *culpa* funda a responsabilidade subjetiva, ao passo que o *risco* constitui no fato de imputação da responsabilidade objetiva. A propósito, em análise ao Código Civil, vemos que o parágrafo único do artigo 927 trata justamente da responsabilidade civil objetiva, senão vejamos:

Art. 927. Aquele que, por ato ilícito (arts. 186 e 187), causar dano a outrem, fica obrigado a repará-lo.

Parágrafo único. Haverá obrigação de reparar o dano, independentemente de culpa, nos casos especificados em lei, ou quando a atividade normalmente desenvolvida pelo autor do dano implicar, por sua natureza, risco para os direitos de outrem.

Pela leitura do dispositivo, concluímos que a obrigação de reparar o dano independe de prova de culpa nos casos especificados em lei e quando o autor do dano criar um risco maior para terceiros, em razão de sua atividade. Considera-se que, mesmo inexistindo lei que regulamente o fato, pode o juiz aplicar o princípio da responsabilidade objetiva, baseando-se no dispositivo legal em questão, “quando a atividade normalmente desenvolvida pelo autor do dano implicar, por sua natureza, risco para os direitos de outrem”.

Assim, sendo o risco o fundamento da responsabilidade, a obrigação de reparar surge do “simples exercício da atividade que o agente desenvolve em seu interesse e sob seu controle, em função do perigo que dela decorre para terceiros.”²⁷⁰. Na conceituação de atividade perigosa, devem ser analisadas aquelas atividades que contenham em si uma grave probabilidade e uma notável potencialidade danosa, de modo que, na conceituação de atividade perigosa, ingressam, como meios hábeis, diferentes veículos, máquinas, objetos e utensílios, tomados em um sentido dinâmico²⁷¹.

Ao aproximar a discussão à problemática analisada neste trabalho, verificamos

²⁶⁹ *Ibidem*.

²⁷⁰ GONÇALVES, Carlos Roberto. **Direito civil brasileiro, volume 4: responsabilidade civil**. 14. ed. São Paulo: Saraiva Educação, 2019. p. 218.

²⁷¹ *Ibidem*, p. 218-219.

que o artigo 27 do Projeto de Lei n.º 2338, de 2023²⁷² – inserido no Capítulo V da Lei, “Da Responsabilidade Civil” – prevê a possibilidade de o fornecedor ou o operador do sistema de inteligência artificial que cause dano patrimonial, moral, individual ou coletivo repará-lo integralmente, independentemente do grau de autonomia do sistema. O § 1º do mesmo artigo dispõe que, quando se tratar de sistema de inteligência artificial de alto risco ou de risco excessivo, o fornecedor ou o operador respondem objetivamente pelos danos causados, na medida de sua participação no dano. Por sua vez, o § 2º estabelece que, quando não se tratar de sistema de inteligência artificial de alto risco, a culpa do agente causador do dano será presumida, aplicando-se a inversão do ônus da prova em favor da vítima.

A partir da leitura dos dispositivos legais, podemos concluir que se buscou responsabilizar na esfera cível danos praticados por sistemas de inteligência artificial, atribuindo-os ao fornecedor ou ao operador do sistema, independentemente do grau de autonomia do sistema. Em sentido contrário, para responsabilização penal, na linha do que sustentamos, é necessária íntima vinculação da atividade do programador ou do usuário, impossibilitada nos sistemas essencialmente autônomos.

Da mesma forma, percebe-se a opção por responsabilizar objetivamente, nos casos de sistemas de alto risco ou de risco excessivo, e com culpa presumida, quando não se tratar de sistema de IA de alto risco. Denota-se, pois, a notória intenção do legislador em evitar eventual impossibilidade de responsabilização pelo dano causado, sendo o dever de reparar, em uma primeira leitura, quase que automático. Doravante, o artigo 28 do PL estabelece hipóteses de não responsabilização:

Art. 28. Os agentes de inteligência artificial não serão responsabilizados quando:

I – comprovarem que não colocaram em circulação, empregaram ou tiraram proveito do sistema de inteligência artificial; ou

II – comprovarem que o dano é decorrente de fato exclusivo da vítima ou de terceiro, assim como de caso fortuito externo.

²⁷² Art. 27. O fornecedor ou operador de sistema de inteligência artificial que cause dano patrimonial, moral, individual ou coletivo é obrigado a repará-lo integralmente, independentemente do grau de autonomia do sistema. § 1º Quando se tratar de sistema de inteligência artificial de alto risco ou de risco excessivo, o fornecedor ou operador respondem objetivamente pelos danos causados, na medida de sua participação no dano. § 2º Quando não se tratar de sistema de inteligência artificial de alto risco, a culpa do agente causador do dano será presumida, aplicando-se a inversão do ônus da prova em favor da vítima.

Insta trazer à baila, ademais, o artigo 36 do PL, que prevê a possibilidade de atribuição de sanções administrativas aplicáveis pela autoridade competente em razão das infrações cometidas às normas previstas na Lei, indicando a natureza extrapenal dos ilícitos previstos. Nesse sentido, dispõe o artigo:

Art. 36. Os agentes de inteligência artificial, em razão das infrações cometidas às normas previstas nesta Lei, ficam sujeitos às seguintes sanções administrativas aplicáveis pela autoridade competente:

I – advertência;

II – multa simples, limitada, no total, a R\$ 50.000.000,00 (cinquenta milhões de reais) por infração, sendo, no caso de pessoa jurídica de direito privado, de até 2% (dois por cento) de seu faturamento, de seu grupo ou conglomerado no Brasil no seu último exercício, excluídos os tributos;

III – publicização da infração após devidamente apurada e confirmada a sua ocorrência;

IV – proibição ou restrição para participar de regime de sandbox regulatório previsto nesta Lei, por até cinco anos;

V – suspensão parcial ou total, temporária ou definitiva, do desenvolvimento, fornecimento ou operação do sistema de inteligência artificial; e

VI – proibição de tratamento de determinadas bases de dados.

Interessante, ainda, a previsão do § 3º, o qual estabelece que “O disposto neste artigo não substitui a aplicação de sanções administrativas, civis ou penais definidas na Lei nº 8.078, de 11 de setembro de 1990, na Lei nº 13.709, de 14 de agosto de 2018, e em legislação específica.”. Com efeito, a Lei nº 8.078/90 corresponde ao chamado “Código de Defesa do Consumidor”, enquanto a Lei nº 13.709/18 é conhecida como “Lei Geral de Proteção de Dados”.

Nesse sentido, civilmente, entendemos que eventual ilícito cometido por um sistema de IA poderá ensejar, além das disposições do Código Civil, na incidência do artigo 12 do Código de Defesa do Consumidor, que corresponde à hipótese de responsabilidade civil objetiva:

Art. 12. O fabricante, o produtor, o construtor, nacional ou estrangeiro, e o importador respondem, independentemente da existência de culpa, pela reparação dos danos causados aos consumidores por defeitos decorrentes de projeto, fabricação, construção, montagem, fórmulas, manipulação, apresentação ou acondicionamento de seus produtos, bem como por informações insuficientes ou inadequadas sobre sua utilização e riscos.

De outra banda, ao analisar a redação do § 3º, verificamos que, no Código de Defesa do Consumidor, estão previstos crimes e disposições de natureza penal no

“Título II”, artigos 61 a 80, do diploma legal. Mantendo coerência com a exposição sistemática construída neste trabalho, compreendemos que as sanções penais referidas somente poderão ser aplicadas na hipótese de o ilícito praticado guardar íntima vinculação à atividade do programador ou do usuário.

Possivelmente, a melhor alternativa extrapenal está no desenvolvimento de medidas preventivas, na linha do que sustenta Carlos Casabona²⁷³. Antes de se pensar em sanções – sejam elas de natureza cível, administrativo ou penal – é impositivo construir métodos para reduzir e dirimir a prática de ilícitos, visando a controlar sistemas de IA que possam oferecer riscos à sociedade.

5 CONCLUSÃO

Em síntese, a partir da análise do que restou exposto, podemos sintetizar e sistematizar as conclusões obtidas por meio dos seguintes pontos:

1. Sistemas de inteligência artificial são tecnologias capazes de realizar funções tradicionalmente feitas por humanos que exigiriam inteligência, simulando o comportamento humano e o processo cognitivo em um computador. Embora a existência da inteligência artificial não seja recente, os últimos anos levaram ao surgimento de novos contornos, em especial pelo desenvolvimento das tecnologias de *machine learning*, cujas máquinas passaram a ter aptidão de adquirir aprendizado a partir de suas próprias experiências. Observa-se, assim, a existência de tecnologias de inteligência artificial autônomas, capazes de tomar decisões sem intervenção humana.

2. Ao passo que o uso da inteligência artificial proporcionou diversos benefícios, especialmente na organização de dados, no auxílio à tomada de decisões e na automação de decisões, coexistem diversos riscos, a exemplo da tomada de decisões tendenciosos, dos riscos de segurança e da contribuição para desinformação e manipulação de informações. Passa-se a surgir, portanto, uma nova criminalidade, na medida em que os sistemas de inteligência artificial podem ser empregados como ferramentas ou agentes para prática de crimes.

²⁷³ PUCRS. **Inteligência artificial pode ter responsabilidade penal**. 2018. Disponível em: <https://www.pucrs.br/blog/inteligencia-artificial-pode-ter-responsabilidade-penal>. Acesso em: 07 ago. 2023.

3. Os riscos da inteligência artificial levaram a um crescente aumento da preocupação dos Estados em regulamentar a inteligência artificial, visando a assegurar um uso seguro, responsável, transparente, auditável, em respeito à privacidade de dados e aos direitos humanos e evitando a utilização para fins indevidos e ilícitos. Nesse contexto, a pioneira Lei de Inteligência Artificial da União Europeia, aprovada em junho de 2023 pelo Parlamento Europeu, e o Projeto de Lei n.º 2.388/2023, em trâmite no Senado Federal brasileiro, vêm ao encontro da necessidade de regulamentação. Notório, com o PL, o surgimento de uma sistemática de supervisão e fiscalização, bem como de responsabilização e aplicação de sanções.

4. O ChatGPT, lançado em novembro de 2022, consiste em uma ferramenta capaz de interagir de forma conversacional, de modo a permitir respostas a perguntas complementares, admitir erros, contestar premissas incorretas e rejeitar solicitações inadequadas. Apesar das diversas possibilidades de utilização e do avanço que representa no cenário tecnológico, ultimamente o ChatGPT vem despertando preocupação, em razão de suas limitações e riscos, dentre eles a possibilidade de ser fonte de desinformação e de preconceito.

5. Muito embora no ordenamento jurídico brasileiro não haja, até o momento, um tipo penal específico que disponha acerca da divulgação de *fake news*, mas tão somente Projetos de Lei, pode-se cogitar a incidência de tipos penais do Código Penal ou do Código Eleitoral, por exemplo, tendo em vista o atingimento a bens jurídicos já existentes. Salutar a existência de um tipo penal específico que trate da conduta de divulgação de notícia falsa, haja vista que a divulgação de *fake news* pode alcançar uma grande dimensão, produzindo resultados altamente danosos, de modo que se estaria tutelando, em verdade, o próprio Estado Democrático de Direito.

6. O ciberespaço se tornou campo para o cometimento de delitos, levando ao surgimento de “crimes digitais”. Nesse cenário, a facilidade em se manifestar tornou propício o discurso e a prática de crimes de ódio, caracterizados pela discriminação e pelo preconceito. No ordenamento jurídico brasileiro, destacam-se as previsões dos incisos XLI e XLII do artigo 5º da CF/88, bem como a Lei n.º 7.716/89, que preveem a necessidade de punição pela prática de discriminação e racismo, seja por raça, cor, etnia, religião ou procedência e, a partir da ADO n.º 26, pela orientação sexual ou identidade de gênero.

7. Os sistemas de inteligência artificial, a exemplo do ChatGPT, são capazes de divulgar *fake news* em larga escala, com maiores frequência e facilidade, contribuindo, inclusive, para o processo de geração de informações falsas mais críveis e efetivas em desinformar. Da mesma forma, a inteligência artificial, treinada com base em dados históricos que podem conter preconceitos e estereótipos, pode perpetuar o racismo e tomar decisões discriminatórias. Denota-se, pois, que os sistemas de IA podem divulgar *fake news* e proferir discurso de ódio assim como os seres humanos.

8. Considerando a possibilidade de os sistemas de inteligência artificial cometerem ilícitos penais, procedeu-se à sistematização em três modelos as hipóteses de responsabilização penal²⁷⁴, associados à maior ou à menor independência que possuem as tecnologias de inteligência artificial frente ao programador ou ao usuário. Aproximando-os ao direito brasileiro, tem-se que, no primeiro modelo, em que os sistemas de inteligência artificial são usados como meros instrumentos para prática de um delito, responderá o programador ou o usuário, tendo em vista a relação de causalidade entre ação – programação ou uso do sistema de IA – e resultado – conduta criminosa produzida. No segundo, os programadores ou usuários, embora possam desconhecer da conduta criminosa cometida por meio de um sistema de IA, serão responsabilizados a título de culpa, por negligência, na hipótese de poderem prever e prevenir o cometimento de uma ofensa. Por sua vez, no terceiro modelo, analisa-se a hipótese de a ação não guardar nenhuma relação de dependência com o programador ou com o usuário.

9. Nos casos de divulgação de *fake news* e de prática de crimes de ódio por sistemas de IA, entendemos que: em primeiro lugar, um sujeito pode desenvolver ou utilizar das aptidões do sistema de IA para, como instrumento, proferir discurso de ódio ou divulgar *fake news*, de modo que, nesta hipótese, responderia penalmente o próprio agente. Em segundo lugar, cogita-se a responsabilização a título de culpa, por negligência, na hipótese de o programador ou o desenvolvedor terem previsto, porém não terem impedido, a prática de ofensas de tais natureza.

10. Usualmente, contudo, diversos casos de prática de crimes de ódio e de

²⁷⁴ HALLEVY, Gabriel. **The Criminal Liability of Artificial Intelligence Entities - from Science Fiction to Legal Social Control**. Akron Intellectual Property Journal: v. 4, n. 2, 2010. Disponível em: <http://doi.org/10.2139/ssrn.1564096>. Acesso em: 09 ago. 2023.

divulgação de *fake news* não possuem relação com a atividade do programador ou do usuário, tendo os sistemas de IA realizado a conduta espontânea e originalmente. A hipótese de produção de resultado típico sem dependência com o programador ou o usuário consiste, justamente, no ponto sensível e nebuloso da matéria.

11. O direito penal brasileiro adota o conceito tripartido e analítico/dogmático de crime, o qual corresponde à ação ou à omissão típica, ilícita e culpável. Da mesma forma, adota a sistemática do crime da teoria finalista, a qual compreende que a ação humana deve ser dirigida a um fim, sendo um acontecimento final, e não causal.

12. Para caracterização da conduta, segundo o direito penal brasileiro, são necessários dois elementos: a ação humana e a vontade dirigida a uma finalidade. Ainda que um sistema de inteligência artificial possa proferir autônoma e independentemente discurso de ódio ou divulgar informações falsas, tais fatos estão associados à própria estrutura do sistema de IA, que “aprende” conforme a interação dos usuários e com as informações constantes de um banco de dados provenientes da internet. Nesse sentido, a conjugação dos dois elementos leva à conclusão de que não é possível a responsabilização penal dos sistemas de IA.

13. A relação de causalidade, prevista no artigo 13 do Código Penal, corresponde a um vínculo material entre a ação humana como causa e determinado evento como efeito. Embora o Código Penal tenha adotado a teoria da equivalência dos antecedentes, são necessárias restrições para correção de uma amplitude excessiva de causa. Nesse sentido, está a teoria da causalidade adequada, que considera que a condição é causa quando se apresenta geralmente proporcionada ou adequada ao resultado, constatado por meio de um juízo de probabilidade. Da mesma forma, a imputação objetiva corresponde a um critério normativo-restritivo da causalidade natural, considerando que a causação de um resultado típico somente realizará o tipo objetivo do delito na hipótese de o agente ter criado um perigo juridicamente desaprovado que levou ao resultado. À conjugação de tais teorias, entendemos que não se faz possível a responsabilização da pessoa física – especialmente o programador – por ilícitos penais praticados pelo sistema de IA, seja com base na teoria da causalidade adequada, visto que não haveria causa relevante para geração do resultado, seja com base na teoria da imputação objetiva, na medida em que, sendo os riscos dos sistemas de IA permitidos, não haverá imputação.

14. Tem-se por impossível a responsabilização da empresa desenvolvedora do sistema de inteligência artificial que comete um ilícito penal, seja porque seriam incapazes de ação, dolo, imprudência, omissão ou culpabilidade, seja porque, até o momento, há apenas duas hipóteses de responsabilização penal da pessoa jurídica existentes do ordenamento jurídico brasileiro: a primeira, prevista no artigo 173, § 5º, da CF/88, ainda não regulamentado, e a segunda, constante no artigo 225, § 3º, da CF/88, regulamentado infraconstitucionalmente pela Lei n.º 9.605/98. Da mesma forma, caso se cogite eventual responsabilização das pessoas jurídicas, deve-se associá-la, necessariamente, à “delinquência empresarial”, não sendo possível se cogitar eventual responsabilização das pessoas jurídicas por crimes comuns, não inseridos no contexto da atividade empresarial.

15. Em que pese a impossibilidade de responsabilização penal, é possível cogitar a incidência da responsabilidade civil, especialmente porque existente a responsabilização objetiva, independentemente de culpa, fundamentada no risco. O artigo 27 do Projeto de Lei n.º 2338, de 2023, nessa linha, prevê a possibilidade de o fornecedor ou o operador do sistema de inteligência artificial que cause dano patrimonial, moral, individual ou coletivo responder civilmente, independentemente do grau de autonomia do sistema. O § 1º do mesmo artigo dispõe que, quando se tratar de sistema de inteligência artificial de alto risco ou de risco excessivo, o fornecedor ou o operador respondem objetivamente pelos danos causados, na medida da sua participação no dano. Por sua vez, o § 2º estabelece que, quando não se tratar de sistema de inteligência artificial de alto risco, a culpa do agente causador do dano será presumida, aplicando-se a inversão do ônus da prova em favor da vítima.

16. O artigo 36 do Projeto de Lei n.º 2338, de 2023, prevê a possibilidade de atribuição de sanções administrativas aplicáveis pela autoridade competente em razão das infrações cometidas às normas previstas na Lei, indicando a natureza extrapenal dos ilícitos previstos. Por seu turno, embora o § 3º preveja a possibilidade de aplicação de sanções penais previstas no Código de Defesa do Consumidor, estas somente poderão ser aplicadas na hipótese de o ilícito praticado guardar íntima vinculação à atividade do programador ou do usuário. Podem, ainda, ser aplicadas sanções de natureza cível, incidindo, além das disposições do Código Civil, o artigo 12 do Código de Defesa do Consumidor.

REFERÊNCIAS

- ABRUSIO, Juliana; Medeiros, Thamara. Os limites da criminalização da desinformação. *In: (Coord.) RAIS, Diogo. Fake News: a conexão entre a desinformação e o direito*. 3. ed. São Paulo: Thomson Reuters Brasil, 2022. *E-book*. Disponível em: <https://proview.thomsonreuters.com/launchapp/title/rt/monografias/156970600>. Acesso em: 23 jul. 2023.
- ALLCOTT, Hunt; GENTZKOW, Matthew. **Social Media and Fake News in the 2016 Election**. *Journal Of Economic Perspectives*, [S.L.], v. 31, n. 2, p. 211-236, 2017. American Economic Association. Disponível em: <http://doi.org/10.1257/jep.31.2.211>. Acesso em: 23 jul. 2023.
- ALVES, Marco Antônio Souza; MACIEL, Emanuella Ribeiro Halfeld. **O fenômeno das fake news: definição, combate e contexto**. *Revista Internet & Sociedade*, [S.L.], v. 1, n. 1, p. 144–171, 2020. Disponível em: <https://revista.internetlab.org.br/o-fenomeno-das-fake-news-definicao-combate-e-contexto>. Acesso em: 22 jul. 2023.
- ARAÚJO FILHO, Marco Antônio Mendes de. **A influência dos elementos de imputação criminal do *Common Law* na interpretação do tipo penal econômico: uma análise comparativa dos crimes contra o mercado de capitais**. *Faculdade Damas, Recife*, v. 7, n. 3, 2021. Disponível em: <http://54.94.8.198/index.php/academico/article/view/1735>. Acesso em: 01 ago. 2023.
- ASSMANN, Hugo. **A metamorfose do aprender na sociedade da informação**. *Ciência da Informação, Brasília*, v. 29, n. 2, p. 07-15, 2000. IBICT. Disponível em: <http://doi.org/10.1590/s0100-19652000000200002>. Acesso em: 23 jul. 2023.
- BARCAROLLO, Felipe. **Inteligência artificial: aspectos ético-jurídicos**. São Paulo: Almedina, 2021. *E-book*. Disponível em: <https://app.minhabiblioteca.com.br/#/books/9786556272801>. Acesso em: 05 jul. 2023.
- BARRETO, Irineu. **Fake News: anatomia da desinformação, discurso de ódio e erosão da democracia**. São Paulo: ExpressaJur, 2022. *E-book*. Disponível em: <https://app.minhabiblioteca.com.br/reader/books/9786555598841>. Acesso em: 28 jul. 2023.
- BBC NEWS BRASIL. **'Fake News' é eleita palavra do ano e ganhará menção em dicionário britânico**. 2017. Disponível em: <https://www.bbc.com/portuguese/internacional-41843695>. Acesso em: 23 jul. 2023.
- BITENCOURT, Cezar Roberto. **Tratado de direito penal: parte especial: crimes contra a pessoa (arts. 121 a 154-B). v. 2**. 23. ed. São Paulo: SaraivaJur, 2023. *E-book*. Disponível em: <https://app.minhabiblioteca.com.br/#/books/9786553627031>. Acesso em: 02 ago. 2023.
- _____. **Tratado de direito penal: parte geral (arts. 1º a 120). v. 1**. 29. ed. São

Paulo: SaraivaJur, 2023. *E-book*. Disponível em:
<https://app.minhabiblioteca.com.br/#/books/9786553627109>. Acesso em: 03 ago. 2023.

BITTAR, Eduardo Carlos Bianca. O Direito na Era Digital: responsabilidade civil e penal pelo uso indevido das redes sociais. *In*: SARLET, Ingo Wolfgang; SARLET, Gabrielle Bezerra Sales; BITTAR, Eduardo Carlos Bianca. **Inteligência artificial, proteção de dados pessoais e responsabilidade na era digital**. São Paulo: ExpressaJur, 2022. *E-book*. Disponível em:
<https://app.minhabiblioteca.com.br/#/books/9786555599527>. Acesso em: 25 jul. 2023.

BRASIL. **Constituição da República Federativa do Brasil de 1988**. Disponível em:
https://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Acesso em: 23 jul. 2023.

_____. **Decreto-Lei nº 2.848, de 7 de dezembro de 1940**. Disponível em:
http://www.planalto.gov.br/ccivil_03/decreto-lei/del2848compilado.htm. Acesso em: 23 jul. 2023.

_____. **Decreto-Lei nº 3.689, de 3 de outubro de 1941**. Disponível em:
http://www.planalto.gov.br/ccivil_03/decreto-lei/del3689compilado.htm. Acesso em: 23 jul. 2023.

_____. **Lei nº 7.716, de 5 de janeiro de 1989**. Disponível em:
https://www.planalto.gov.br/ccivil_03/leis/l7716.htm. Acesso em: 23 jul. 2023.

_____. **Lei nº 8.078, de 11 de setembro de 1990**. Disponível em:
http://www.planalto.gov.br/ccivil_03/leis/l8078.htm. Acesso em: 23 jul. 2023.

_____. **Lei nº 10.406, de 10 de janeiro de 2002**. Disponível em:
http://www.planalto.gov.br/ccivil_03/LEIS/2002/L10406.htm. Acesso em: 23 jul. 2023.

_____. **Lei nº 13.709, de 14 de agosto de 2018**. Disponível em:
https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm. Acesso em: 23 jul. 2023.

_____. **Lei nº 14.532, de 11 de agosto de 2023**. Disponível em:
https://www.planalto.gov.br/ccivil_03/_ato2023-2026/2023/lei/l14532.htm. Acesso em: 23 jul. 2023.

_____. Supremo Tribunal Federal. **Ação Direta de Inconstitucionalidade por Omissão n.º 26**. Requerente: Partido Popular Socialista. Relator: Ministro Celso de Mello. Tribunal Pleno. Julgado em 13/06/2019. Disponível em:
<https://redir.stf.jus.br/paginadorpub/paginador.jsp?docTP=TP&docID=754019240>. Acesso em: 25 jul. 2023.

_____. Supremo Tribunal Federal. **Recurso Extraordinário n.º 548.181**. Recorrente: Ministério Público Federal. Recorrido: Petróleo Brasileiro S/A – Petrobrás. Relatora: Ministra Rosa Weber. Primeira Turma. Julgado em 06/08/2013. Disponível em:
<https://redir.stf.jus.br/paginadorpub/paginador.jsp?docTP=TP&docID=7087018>. Acesso em: 06 ago. 2023.

BRUGGER, Winfried. **Proibição ou proteção do discurso do ódio?: algumas observações sobre o direito alemão e o americano**. Direito Público, Porto Alegre, v. 4, n. 15, 2007. Disponível em: <https://www.portaldeperiodicos.idp.edu.br/direitopublico/article/view/1418>. Acesso em: jul. 2023.

CALDWELL, Matthew; ANDREWS, Jerone. T. A.; TANAY, Thomas; GRIFFIN, Lewis. D.. **AI-enabled future crime**. Crime Science, [S.L.], v. 9, n. 1, 2020. Springer Science and Business Media LLC. Disponível em: <https://doi.org/10.1186/s40163-020-00123-8>. Acesso em: 05 jul. 2023.

CÂMARA DOS DEPUTADOS. **PL 6812/2017**. Disponível em: <https://www.camara.leg.br/proposicoesWeb/fichadetramitacao?idProposicao=2122678>. Acesso em: 23 jul. 2023.

CAMBRIDGE ENGLISH DICTIONARY. **Fake news**. Disponível em: <https://dictionary.cambridge.org/us/dictionary/english/fake-news>. Acesso em: 22 jul. 2023.

CAVALIERI FILHO, Sérgio. **Programa de responsabilidade civil**. 16. ed. Barueri: Atlas, 2023. *E-book*. Disponível em: <https://app.minhabiblioteca.com.br/#/books/9786559775217>. Acesso em: 07 ago. 2023.

COLLINS ENGLISH DICTIONARY. **Definition of 'fake news'**. Disponível em: <https://www.collinsdictionary.com/dictionary/english/fake-news>. Acesso em: 22 jul. 2023.

COMISSÃO EUROPEIA. **Regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial (regulamento inteligência artificial) e altera determinados atos legislativos da União**. Bruxelas, 2021. Disponível em: <https://eur-lex.europa.eu/legal-content/PT/TXT/?uri=CELEX:52021PC0206>. Acesso em: 04 jul. 2023.

CONSULTOR JURÍDICO. **Publicações de cunho nazista no Twitter configuram crime de racismo, diz juiz**. 2023. Disponível em: <https://www.conjur.com.br/2023-jul-17/publicacoes-cunho-nazista-twitter-configuram-crime-racismo>. Acesso em: 25 jul. 2023.

COPPIN, Ben. **Inteligência artificial**. Rio de Janeiro: LTC, 2013. Tradução de: Jorge Duarte Pires. *E-book*. Disponível em: <https://app.minhabiblioteca.com.br/#/books/978-85-216-2936-8>. Acesso em: 04 jul. 2023.

CRESPO, Marcos Xavier de Freitas. **Crimes digitais**. São Paulo: Saraiva, 2011. *E-book*. Disponível em: <https://app.minhabiblioteca.com.br/#/books/9788502136663>. Acesso em: 25 jul. 2023.

DIAS, José de Aguiar. **Da responsabilidade civil**. Rio de Janeiro: Forense, 1944. p. 11.

DIETTERICH, Thomas G.; HORVITZ, Eric J. **Rise of concerns about**

AI. *Communications Of The Acm*, v. 58, n. 10, 2015. Disponível em: <https://doi.org/10.1145/2770869>. Acesso em: 05 jul. 2023.

DOTTI, Rene Ariel. **Curso de direito penal: parte geral**. Rio de Janeiro: Forense, 2001.

DUARTE, Filipe Ribeiro. **ChatGPT 4: os pontos de atenção e os riscos legais trazidos pelo uso da ferramenta**. 2023. Disponível em: <https://lawinnovation.com.br/chatgpt-4-os-pontos-de-atencao-e-os-riscos-legais-trazidos-pelo-uso-da-ferramenta>. Acesso em: 05 jul. 2023.

DUARTE, Maria Carolina de Almeida. **Globalização e a Nova Criminalidade**. *Revista Territórios e Fronteiras*, Cuiabá, v. 2, n. 1, 2011. *Revista Territórios e Fronteiras*. Disponível em: <https://periodicoscientificos.ufmt.br/territoriosefronteiras/index.php/v03n02/article/view/32/31>. Acesso em: 05 jul. 2023.

ESPINOSA, Alejandra Morán. **Responsabilidad penal de la Inteligencia Artificial (IA). ¿La próxima frontera?** *Revista del Instituto de Ciencias Jurídicas de Puebla*, México, Puebla, México, v. 15, n. 48, 2021. Disponível em: <https://doi.org/10.35487/rius.v15i48.2021.706>. Acesso em: 04 jul. 2023.

ESTELLITA, Heloisa. **A responsabilidade isolada da pessoa jurídica por crimes ambientais**. 2014. Disponível em: <https://www.conjur.com.br/2014-dez-16/heloisa-estellita-responsabilidade-isolada-crimes-ambientais#author>. Acesso em: 06 ago. 2023.

FACELI, Katti; LORENA, Ana Carolina; GAMA, João; ALMEIDA, Tiago Agostinho de; CARVALHO, André C. P. L. F. de. **Inteligência artificial: uma abordagem de abordagem de aprendizado de máquina**. 2. ed. Rio de Janeiro: LTC, 2021. *E-book*. Disponível em: <https://app.minhabiblioteca.com.br/#/books/9788521637509>. Acesso em: 05 jul. 2023.

FAHIM, Sadaf. **AI and Criminal Liability**. *Indian Journal of Artificial Intelligence and Law*: v. 1, n. 1. 2020, Disponível em: <http://doi.org/10.2139/ssrn.1564096>. Acesso em: 01 ago. 2023.

FOLHA DE SÃO PAULO. **ChatGPT será maior espalhador de desinformação que já existiu, diz pesquisador**. 2023. Disponível em: <https://www1.folha.uol.com.br/tec/2023/02/chatgpt-sera-maior-espalhador-de-desinformacao-que-ja-existiu-diz-pesquisador.shtml>. Acesso em: 26 jul. 2023.

_____. **Conservadores buscam versão do ChatGPT que dê respostas com visão da direita**. 2023. Disponível em: <https://www1.folha.uol.com.br/tec/2023/03/conservadores-buscam-versao-do-chatgpt-que-de-respostas-com-visao-da-direita.shtml>. Acesso em: 27 jul. 2023.

FRAGOSO, Heleno Cláudio. **Lições de Direito Penal - A Nova Parte Geral**. 7. ed. Rio de Janeiro: Forense, 1985.

FREITAS, Pedro Miguel; ANDRADE, Francisco; NOVAIS, Paulo. **Criminal Liability of Autonomous Agents: From the Unthinkable to the Plausible**. *In*: (Coord.)

CASANOVA, Pompeus; PAGALLO, Ugo; PALMIRANI, Monica; SARTOR, Giovanni. *AI Approaches to the Complexity of Legal Systems*. Berlin, Heidelberg: Springer, 2014. Disponível em: https://doi.org/10.1007/978-3-662-45960-7_11. Acesso em: 01 ago. 2023.

G1 GLOBO. **China registra primeira prisão por notícias falsas geradas pelo ChatGPT**. 2023. Disponível em: <https://g1.globo.com/tecnologia/noticia/2023/05/10/china-registra-primeira-prisao-por-noticias-falsas-geradas-pelo-chatgpt.ghtml>. Acesso em: 27 jul. 2023.

_____. **Terrorismo em Brasília: o dia em que bolsonaristas criminosos depredaram Planalto, Congresso e STF**. 2023. Disponível em: <https://g1.globo.com/df/distrito-federal/noticia/2023/01/08/o-dia-em-que-bolsonaristas-invalidaram-o-congresso-o-planalto-e-o-stf-como-isso-aconteceu-e-quais-as-consequencias.ghtml>. Acesso em: 24 jul. 2023.

GELFERT, Axel. **Fake News: a definition**. *Informal Logic*, [S.L.], v. 38, n. 1, p. 84-117, 2018. Disponível em: <http://doi.org/10.22329/il.v38i1.5068>. Acesso em: 22 jul. 2023.

GIANSCIRACUSA, Noah. **How Algorithms Create and Prevent Fake News**. Nova York: Springer Science Business Medida, 2021. *E-book*. Disponível em: <https://doi.org/10.1007/978-1-4842-7155-1> Acesso em: 25 jul. 2023.

GIUSTI, Serena; PIRAS, Elisa. **Democracy and Fake News: Information Manipulation and Post-Truth Politics**. Routledge, [S.L.], 2021. *E-book*. Disponível em: <https://search.ebscohost.com/login.aspx?direct=true&AuthType=shib&db=nlebk&AN=2681917&lang=pt-br&scope=site&authtype=guest,shib&custid=s5837110&groupid=main&profile=eds>. Acesso em: 22 jul. 2023.

GIZMODO. **Bing's AI Prompted a User to Say 'Heil Hitler'**. 2023. Disponível em: <https://gizmodo.com/ai-bing-microsoft-chatgpt-heil-hitler-prompt-google-1850109362>. Acesso em: 27 jul. 2023.

GONÇALVES, Carlos Roberto. **Direito civil brasileiro, volume 4: responsabilidade civil**. 14. ed. São Paulo: Saraiva Educação, 2019.

GOOGLE SUPPORT. **Política de discurso de ódio**. Disponível em: <https://support.google.com/youtube/answer/2801939?hl=pt-BR>. 2023. Acesso em: 25 jul. 2023.

HALLEVY, Gabriel. **The Criminal Liability of Artificial Intelligence Entities - from Science Fiction to Legal Social Control**. *Akron Intellectual Property Journal*: v. 4, n. 2, 2010. Disponível em: <http://doi.org/10.2139/ssrn.1564096>. Acesso em: 04 jul. 2023.

HARIRI, Walid. **Unlocking the Potential of ChatGPT: a comprehensive exploration of its applications, advantages, limitations, and future directions in natural language processing**. Annaba, Argélia, 2023. Disponível em: <https://doi.org/10.48550/arXiv.2304.02017>. Acesso em: 06 jul. 2023.

HILGENDORF, Eric; VALERIUS, Brian. **Direito Penal: parte geral**. Tradução de: Orlandino Gleizer. São Paulo: Marcial Pons, 2019.

HUNGRIA, Néelson; FRAGOSO, Heleno Cláudio. **Comentários ao Código Penal, volume I, tomo II: arts. 11 ao 27**. 6. ed. Rio de Janeiro: Forense, 1983.

JACOBS, James B; POTTER, Kimberly. **Hate crimes: criminal law & identity politics**. Oxford University Press: Nova York, 2001. *E-book*. Disponível em: <https://search.ebscohost.com/login.aspx?direct=true&AuthType=shib&db=nlebk&AN=146872&lang=pt-br&scope=site&authtype=guest,shib&custid=s5837110&groupid=main&profile=eds>. Acesso em: 25 jul. 2023.

JESUS, Damásio Evangelista de. MILAGRE, José Antonio. **Manual de crimes informáticos**. São Paulo: Saraiva, 2016. *E-book*. <https://app.minhabiblioteca.com.br/#/books/9788502627246>. Acesso em: 02 ago. 2023.

_____. Proteção ambiental e responsabilidade penal da pessoa jurídica. Carta Forense, São Paulo, n. 154, 2016 *apud* NETTO, Alamiro Velludo Salvador. **Responsabilidade penal da pessoa jurídica**. 3. ed. São Paulo: Thomson Reuters Brasil, 2023. *E-book*. Disponível em: <https://proview.thomsonreuters.com/launchapp/title/rt/monografias/148836979>.

JOTA. **Inteligência artificial pode perpetuar o racismo, diz diretora da DataSphere**. 2023. Disponível em: <https://www.jota.info/jotinhas/inteligencia-artificial-pode-perpetuar-o-racismo-diz-diretora-da-datasphere-17042023>. Acesso em: 27 jul. 2023.

KAUFMAN, Dora. **A inteligência artificial irá suplantar a inteligência humana?** Barueri: Estação das Letras e Cores, 2019.

_____. **Desmistificando a inteligência artificial**. Belo Horizonte: Autêntica, 2022. *E-book*. Disponível em: <https://app.minhabiblioteca.com.br/#/books/9786559281596>. Acesso em: 04 jul. 2023.

KINGSTON, J. K. C. Artificial Intelligence and Legal Liability. *In*: (Coord.) BRAMER, Max; PETRIDIS, Miltos. **Research and Development in Intelligent Systems XXXIII**. [S.L.]: Springer, 2016. *E-book*. Disponível em: <https://doi.org/10.1007/978-3-319-47175-4>. Acesso em: 01 ago. 2023.

LILLEKER, Darren G.. **Evidence to the Culture, Media and Sport Committee. 'Fake news' inquiry presented by the Faculty for Media & Communication**, Bournemouth University, 2017. Disponível em: <https://eprints.bournemouth.ac.uk/28610/3/Evidence%20Submission%20-%20Fake%20News%20FINAL.pdf>. Acesso em: 23 jul. 2023.

MARTINS-COSTA, Judith. A linguagem da responsabilidade civil. *In*: (Coord.) BIANCHI, José Flávio; MENDONÇA PINHEIRO, Rodrigo Gomes de; ARRUDA ALVIM, Teresa. **Jurisdição e Direito Privado: Estudos em homenagem aos 20 anos da Ministra Nancy Andrighi no STJ**. São Paulo: Revista dos Tribunais, 2020. *E-book*. Disponível em:

<https://proview.thomsonreuters.com/launchapp/title/rt/monografias/230263461>. Acesso em: 07 ago. 2023.

MELLO, Patrícia Campos. **A máquina do ódio: notas de uma repórter sobre fake news e violência digital**. Companhia das Letras, São Paulo, 2020.

MELO, Ana Karolina Acris; SOUZA, Gêssica Cardoso; VASCO, Amanda Corrêa; REIS, Bruno Salomon. **Regulação da inteligência artificial: benchmarking de países selecionados**. [S.L.]: Escola Nacional de Administração Pública, 2022. Disponível em: <http://repositorio.enap.gov.br/handle/1/7419>. Acesso em: 04 jul. 2023.

NETTO, Alamiro Velludo Salvador. **Responsabilidade penal da pessoa jurídica**. 3. ed. São Paulo: Thomson Reuters Brasil, 2023. E-book. Disponível em: <https://proview.thomsonreuters.com/launchapp/title/rt/monografias/148836979>. Acesso em: 06 ago. 2023.

NORVIG, Peter; RUSSEL, Stuart Jonathan. **Inteligência Artificial**. Rio de Janeiro: LTC, 2013. Tradução de Regina Célia Simille de Macedo. *E-book*. Disponível em: <https://app.minhabiblioteca.com.br/#/books/9788595156104>. Acesso em: 05 jul. 2023.

_____. **Inteligência artificial: uma abordagem moderna**. 4. Ed. Rio de Janeiro: LTC, 2022. Tradução de: Daniel Vieira e Flávio Soares Corrêa da Silva. *E-book*. Disponível em: <https://app.minhabiblioteca.com.br/#/books/9788595159495>. Acesso em: 04 jul. 2023.

NUCCI, Guilherme de Souza. **Curso de direito penal: parte geral: arts. 1º a 120 do Código Penal**. 7. ed. Rio de Janeiro: Forense, 2023. *E-book*. Disponível em: <https://app.minhabiblioteca.com.br/#/books/9786559646852>. Acesso em: 03 ago. 2023.

OFFICIAL MICROSOFT BLOG. **Learning from Tay's introduction**. 2016. Disponível em: <https://blogs.microsoft.com/blog/2016/03/25/learning-tays-introduction/>. Acesso em: 27 jul. 2023.

OPENAI. **Introducing ChatGPT**. 2022. Disponível em: <https://openai.com/blog/chatgpt>. Acesso em: 05 jul. 2023.

PARLAMENTO EUROPEU. **Lei da UE sobre IA: primeira regulamentação de inteligência artificial**. 2023. Disponível em: <https://www.europarl.europa.eu/news/pt/headlines/society/20230601STO93804/lei-da-ue-sobre-ia-primeira-regulamentacao-de-inteligencia-artificial>. Acesso em: 04 jul. 2023.

_____. **Parlamento negocia primeiras regras para inteligência artificial mais segura**. 2023. Disponível em: <https://www.europarl.europa.eu/news/pt/press-room/20230609IPR96212/parlamento-negoceia-primeiras-regras-para-inteligencia-artificial-mais-segura>. Acesso em: 04 jul. 2023.

PEIXOTO, Fabiano Hartmann; SILVA, Roberta Zumblick Martins da. **Inteligência artificial e direito**. Curitiba: Alteridade Editora, 2019.

PERRY, Barbara. **In the Name of Hate: Understanding Hate Crimes**. Routledge:

Nova York, 2001. *E-book*. Disponível em: <https://doi.org/10.4324/9780203905135>. Acesso em: 25 jul. 2023.

PONTES, Marcos; ALVIM, Paulo Cesar Rezende de Carvalho; GONTIJO, José Gustavo Sampaio; AZAMBUJA, Eliana Cardoso Emediato de; VIDAL, Karina Domingues Bressan; OLIVEIRA, Cezar Luciano Cavalcanti de. Inteligência artificial no contexto da estratégia brasileira de transformação digital. *In*: (Coord.) VAINZOF, Rony; GUTIERREZ, Andriei. **Inteligência Artificial (IA)**. São Paulo: Revista dos Tribunais, 2021. *E-book*. Disponível em: <https://proview.thomsonreuters.com/launchapp/title/rt/monografias/270029414>. Acesso em: 05 jul. 2023.

PRADO, Luiz Regis. **Curso de direito penal brasileiro: volume 1**. 6. ed. São Paulo: Revista dos Tribunais, 2018. *E-book*. Disponível em: <https://proview.thomsonreuters.com/launchapp/title/rt/monografias/122952706>. Acesso em: 04 ago. 2023.

PUCRS. **Inteligência artificial pode ter responsabilidade penal**. 2018. Disponível em: <https://www.pucrs.br/blog/inteligencia-artificial-pode-ter-responsabilidade-penal>. Acesso em: 03 ago. 2023.

ROBLES-LESSA, Moyana Mariano; CABRAL, Hildeliza Lacerca Tinoco Boechat; SILVESTRE, Gilberto Fachetti. **Deepfake: a inteligência artificial e o algoritmo causando riscos à sociedade no ciberespaço**. Derecho y Cambio Social, Lima-Peru, n. 61, Jul./Set., 2020.

RODRIGUES, Ana Paula da Fonseca; SILVA, Roberto Ferreira Archanjo da; POUSADA; Estevan Lo Ré. **Inteligência artificial, lesão a bens jurídicos penais e responsabilidade penal**. Relações Internacionais do Mundo Atual Unicuritiba: v. 2, n. 35. 2022, Disponível em: <http://revista.unicuritiba.edu.br/index.php/RIMA/article/view/5884/371373858>. Acesso em: 02 ago. 2023.

ROSADO, David. **The Political Biases of ChatGPT**. Hamilton, Nova Zelândia, 2023. Disponível em: <https://doi.org/10.3390/socsci12030148>. Acesso em: 06 jul. 2023.

SANTOS, Christiano Jorge. **Crimes de preconceito e de discriminação**. 2. ed. São Paulo: Saraiva, 2010. *E-book*. Disponível em: <https://app.minhabiblioteca.com.br/#/books/9788502113114>. Acesso em: 25 jul. 2023.

SANTOS, Juarez Cirino dos. **Direito penal: parte geral**. 8. ed. Florianópolis: Tirant lo Blanch, 2018.

SANTOS, Marcelo Henrique dos. **Introdução à inteligência artificial**. Londrina: Editora e Distribuidora Educacional S.A., 2021. *E-book*. Disponível em: <https://app.minhabiblioteca.com.br/#/books/9786559031245>. Acesso em: 04 jul. 2023.

SANTOS, Milton. **Técnica, espaço, tempo: globalização e meio técnico-científico informacional**. São Paulo: Hucitec, 1994. *E-book*. Disponível em: <http://geocrocetti.com/msantos/tecnica.pdf>. Acesso em: 05 jul. 2023.

SCHWAB, Klaus. **A Quarta Revolução Industrial**. São Paulo: Edipro, 2019. *E-book*. Disponível em: <https://books.google.com.br/books?id=XZSWDwAAQBAJ>. Acesso em: 05 jul. 2023.

SENADO FEDERAL. **Projeto de Lei do Senado nº 471, de 2018**. Disponível em: <https://www25.senado.leg.br/web/atividade/materias/-/materia/134781>. Acesso em: 23 jul. 2023.

_____. **Projeto de Lei do Senado nº 473, de 2017**. Disponível em: <https://www25.senado.leg.br/web/atividade/materias/-/materia/131758>. Acesso em: 23 jul. 2023.

_____. **Projeto de Lei do Senado nº 533, de 2018**. Disponível em: <https://www25.senado.leg.br/web/atividade/materias/-/materia/134952>. Acesso em: 23 jul. 2023.

_____. **Projeto de Lei nº 2338, de 2023**. Disponível em: <https://www25.senado.leg.br/web/atividade/materias/-/materia/157233>. Acesso em: 04 jul. 2023.

_____. **Projeto de Lei nº 3813, de 2021**. Disponível em: <https://www25.senado.leg.br/web/atividade/materias/-/materia/150544>. Acesso em: 23 jul. 2023.

_____. **Regimento Interno do Senado Federal**. Disponível em: <https://www25.senado.leg.br/web/atividade/legislacao/regimento-interno>. Acesso em: 23 jul. 2023.

_____. **Uso de inteligência artificial pelo poder público será sujeito a regulamentação**. 2023. Disponível em: <https://www12.senado.leg.br/noticias/materias/2023/05/12/uso-de-inteligencia-artificial-pelo-poder-publico-sera-sujeito-a-regulamentacao>. Acesso em: 04 jul. 2023.

SICHMAN, Jaime Simão. **Inteligência Artificial e sociedade: avanços e riscos**. Estudos Avançados, São Paulo, v. 35, 2021. Disponível em: <https://doi.org/10.1590/s0103-4014.2021.35101.004>. Acesso em: 05 jul. 2023.

SILVA, Ângelo Roberto Ilha da. **Curso de direito penal: parte geral**. Porto Alegre: Livraria do Advogado. 2020.

SILVA, Marco Antônio Marques da; BRAGA, Mariana Stuart Nogueira. **Fake news: é necessário novo tipo penal?**. Revista Paradigma, [S.L.], v. 30, n. 2, p. 152–174, 2022. Disponível em: <https://revistas.unaerp.br/paradigma/article/view/2433>. Acesso em: 22 jul. 2023.

SILVA, Nilton Correia da. Parte I - Compreensão da inteligência artificial e dos seus pressupostos de controle e regulação. *In*: (Coord.) FRAZÃO, Ana; MULHOLLAND, Caitlin. **Inteligência artificial e direito: ética, regulação e responsabilidade**. 2. ed. São Paulo: Thomson Reuters Brasil, 2020. *E-book*. Disponível em: <https://proview.thomsonreuters.com/launchapp/title/rt/monografias/198921448>. Acesso em: 04 jul. 2023.

SILVA, Tarcízio. **Racismo algorítmico: inteligência artificial e discriminação nas redes digitais**. São Paulo: Edições Sesc, 2022. *E-book*. Disponível em: <https://books.google.com.br/books?id=BCZcEAAAQBA>. Acesso em: 05 jul. 2023.

SILVA, Thiago Vitor Lima da; BEZERRA, Germano Cunha. **Responsabilidade penal nos delitos de inteligência artificial**. Revista Eletrônica Nova Roma: v. 2, n. 1, 2018. Disponível em: <https://faculadadenovaroma.com.br/revista/revistas/escola-de-direito-2.pdf#page=80>. Acesso em: 03 ago. 2023.

SOLAIMAN, Irene *et al.* **Release Strategies and the Social Impacts of Language Models**. Arxiv Cornell Universty, 2019. Disponível em: <https://arxiv.org/ftp/arxiv/papers/1908/1908.09203.pdf>. Acesso em: 27 jul. 2023.

SPENCER, Michael K.. **Deep Fake, a mais recente ameaça distópica**. 2019. Tradução de: Gabriela Leite. Disponível em: <https://ihu.unisinos.br/78-noticias/589686-deep-fake-a-mais-recente-ameaca-distopica>. Acesso em: 05 jul. 2023.

SPITALE, Giovanni; BILLER-ANDORNO, Nikola; GERMANI, Federico. **AI model GPT-3 (dis)inform us better than humans**. Science Advances: v. 9, n. 26. 2023. Disponível em: <https://www.science.org/doi/10.1126/sciadv.adh1850>. Acesso em: 04 jul. 2023.

STEIBEL, Fabro; VICENTE, Victor Freitas; JESUS, Diego Santos Vieira de. Possibilidades e potenciais da utilização da Inteligência Artificial. *In:* (Coord.) FRAZÃO, Ana; MULHOLLAND, Caitlin. **Inteligência artificial e direito: ética, regulação e responsabilidade**. 2. ed. São Paulo: Thomson Reuters Brasil, 2020. *E-book*. Disponível em: <https://proview.thomsonreuters.com/launchapp/title/rt/monografias/198921448>. Acesso em: 04 jul. 2023.

TANDOC, Edson C.; LIM, Zheng Wei; LING, Richard. **Defining “Fake News”**. Digital Journalism, [S.L.], v. 6, n. 2, p. 137-153, 2017. Disponível em: <https://doi.org/10.1080/21670811.2017.1360143>. Acesso em: 22 jul. 2023.

TAULLI, Tom. **Introdução à inteligência artificial**. São Paulo: Novatec, 2020. Tradução de: Luciana do Amaral Teixeira.

TAVARES, Juarez. **Teoria do Injusto Penal**. 4. ed. São Paulo: Tirant lo Blanch, 2019.

TEIDER, Lucas Hinckel; SANTOS, Gabriel Pivatto dos. **Responsabilidade penal da inteligência artificial (?): a problemática relacionada ao elemento da conduta na clássica estrutura analítica do delito**. *In:* (Coord.) GONZÁLEZ, Javier Garcia; LOZANO, Álvaro Alzina; RODRÍGUEZ, Gabriel Martín; GONÇALVES, Rubén Miranda; VEIGA, Fábio da Silva. El derecho público y privado ante las nuevas tecnologías. Madrid: Dykinson, 2020. Disponível em: <https://doi.org/10.2307/j.ctv153k4rn>. Acesso em: 04 jul. 2023.

TEPEDINO, Gustavo; SILVA, Rodrigo da Guia. **Desafios da inteligência artificial em matéria de responsabilidade civil**. Revista Brasileira de Direito Civil, Belo Horizonte, v. 21, n. 2, 2019. Disponível em:

<https://rbdcivil.emnuvens.com.br/rbdc/article/view/465>. Acesso em: 04 jul. 2023.

TILT UOL. **EUA investigam ChatGPT para avaliar riscos de fake news e uso indevido de dados**. 2023. Disponível em:

<https://www.uol.com.br/tilt/noticias/afp/2023/07/14/eua-investiga-chatgpt-para-avaliar-riscos-de-fake-news-e-uso-indevido-de-dados.html>. Acesso em: 27 jul. 2023.

VEJA. **Exposto à internet, robô da Microsoft vira racista em 1 dia**. 2016.

Disponível em: <https://veja.abril.com.br/tecnologia/exposto-a-internet-robo-da-microsoft-vira-racista-em-1-dia>. Acesso em: 27 jul. 2023.

VIVIANI, Rodrigo Andrade. **Responsabilidade penal da pessoa jurídica: aspectos controvertidos no direito brasileiro**. Curitiba: Juruá, 2008.

WELZEL, Hans. **Direito Penal**. Tradução de: Afonso Celso Rezende. Campinas: Romana, 2003.

_____. **O novo sistema jurídico-penal: uma introdução à doutrina da ação finalista**. Tradução de: Luiz Regis Prado. São Paulo: Revista dos Tribunais, 2015. *E-book*. Disponível em:

<https://proview.thomsonreuters.com/launchapp/title/rt/monografias/103790584/v4/document/103909793>. Acesso em: 26 ago. 2023.

ZAFFARONI, Eugenio Raúl; BATISTA, Nilo; ALAGIA, Alejandro; SLOKAR, Alejandro. **Direito penal brasileiro, segundo volume: teoria do delito: introdução histórica e metodológica, ação e tipicidade**. 2. ed. Rio de Janeiro: Revan, 2010.

ZAFFARONI, Eugenio Raúl; PIERANGELI, José Henrique. **Manual de direito penal brasileiro: parte geral**. 2. ed. São Paulo: Revista dos Tribunais, 1999.